

SEARCH REQUEST FORM

Scientific and Technical Information Center

Requester's Full Name: _____ Examiner #: _____ Date: _____
 Art Unit: _____ Phone Number 30 _____ Serial Number: _____
 Mail Box and Bldg/Room Location: _____ Results Format Preferred (circle): PAPER DISK E-MAIL

If more than one search is submitted, please prioritize searches in order of need.

 Please provide a detailed statement of the search topic, and describe as specifically as possible the subject matter to be searched. Include the elected species or structures, keywords, synonyms, acronyms, and registry numbers, and combine with the concept or utility of the invention. Define any terms that may have a special meaning. Give examples or relevant citations, authors, etc, if known. Please attach a copy of the cover sheet, pertinent claims, and abstract.

Title of Invention: _____

Inventors (please provide full names): _____

Earliest Priority Filing Date: _____

**For Sequence Searches Only* Please include all pertinent information (parent, child, divisional, or issued patent numbers) along with the appropriate serial number.*

STAFF USE ONLY

Searcher: D. Schreber

Searcher Phone #: 308-4242

Searcher Location: CMI 12E18

Date Searcher Picked Up: _____

Date Completed: 5/1

Searcher Prep & Review Time: 9

Clerical Prep Time: _____

Online Time: 82

Type of Search

NA Sequence (#) 2

AA Sequence (#) 3

Structure (#) _____

Bibliographic _____

Litigation _____

Fulltext _____

Patent Family _____

Other _____

Vendors and cost where applicable

STN: _____

Dialox: _____

Questel/Orbit: _____

Dr Link: _____

Lexis/Nexis: _____

Sequence Systems: Compugen Ig Genet FastDB

WWW/Internet: _____

Other (specify): _____

100

101

102

103

104

105

106

107

108

109

110

111

112

113

114

115

116

117

118

119

120

121

122

123

124

125

126

127

128

129

130

131

132

133

134

135

136

137

138

139

140

141

142

143

144

145

146

147

148

149

150

151

152

153

154

155

156

157

158

159

160

161

162

163

164

165

166

167

168

169

170

171

172

173

174

175

176

177

178

179

180

181

182

183

184

185

186

187

188

189

190

191

192

193

194

195

196

197

198

199

200

201

202

203

204

205

206

207

208

209

210

211

212

213

214

215

216

217

218

219

220

221

222

223

224

225

226

227

228

229

230

231

232

233

234

235

236

237

238

239

240

241

242

243

244

245

246

247

248

249

250

251

252

253

254

255

256

257

258

259

260

261

262

263

264

265

266

267

268

269

270

271

272

273

274

275

276

277

278

279

280

281

282

283

284

285

286

287

288

289

290

291

292

293

294

295

296

297

298

299

300

301

302

303

304

305

306

307

308

309

310

311

312

313

314

315

316

317

318

319

320

321

322

323

324

325

326

327

328

329

330

331

332

333

334

335

336

337

338

339

340

341

342

343

344

345

346

347

348

349

350

351

352

353

354

355

356

357

358

359

360

361

362

363

364

365

366

367

368

369

370

371

372

373

374

375

376

377

378

379

380

381

382

383

384

385

386

387

388

389

390

391

392

393

394

395

396

397

398

399

400

401

402

403

404

405

406

407

408

409

410

411

412

413

414

415

416

417

418

419

420

421

422

423

424

425

426

427

428

429

430

431

432

433

434

435

436

437

438

439

440

441

442

443

444

445

446

447

448

449

450

451

452

453

454

455

456

457

458

459

460

461

462

463

464

465

466

467

468

469

470

471

472

473

474

475

476

477

478

479

480

481

482

483

484

485

486

487

488

489

490

491

492

493

494

495

496

497

498

499

500

501

502

503

504

505

506

507

508

509

510

511

512

513

514

515

516

517

518

519

520

521

522

523

524

525

526

527

528

529

530

531

532

533

534

535

536

537

538

539

540

541

542

543

544

545

546

547

548

549

550

551

552

553

554

555

556

557

558

559

560

561

562

563

564

565

566

567

568

569

570

571

572

573

574

575

576

577

578

579

580

581

582

583

584

585

586

587

588

589

590

591

592

593

594

595

596

597

598

599

600

601

602

603

604

605

606

607

608

609

610

611

612

613

614

615

616

617

618

619

620

621

622

623

624

625

626

627

628

629

630

631

632

633

634

635

636

637

638

639

640

641

642

643

644

645

646

647

648

649

650

651

652

653

654

655

656

657

658

659

660

661

662

663

664

665

666

667

668

669

670

671

672

673

674

675

676

677

678

679

680

681

682

683

684

685

686

687

688

689

690

691

692

693

694

695

696

697

698

699

700

701

702

703

704

705

706

707

708

709

710

711

712

713

714

715

716

717

718

719

720

721

722

723

724

725

726

727

728

729

730

731

732

733

734

735

736

737

738

739

740

741

742

743

744

745

746

747

748

749

750

751

752

753

754

755

756

757

758

759

760

761

762

763

764

765

766

767

768

769

770

771

772

773

774

775

776

777

778

779

780

781

782

783

784

785

786

787

788

789

790

791

792

793

794

795

796

797

798

799

800

801

802

803

804

805

806

807

808

809

810

811

812

813

814

815

816

817

818

819

820

821

822

823

824

825

826

827

828

829

830

831

832

833

834

835

836

837

838

839

840

841

842

843

844

845

846

847

848

849

850

851

852

853

854

855

856

857

858

859

860

861

862

863

864

865

866

867

868

869

870

871

872

873

874

875

876

877

878

879

880

881

882

883

884

885

886

887

888

889

890

891

892

893

894

895

896

897

898

899

900

901

902

903

904

905

906

907

908

909

910

911

912

913

914

915

916

917

918

919

920

921

922

923

924

925

926

927

928

929

930

931

932

933

934

935

936

937

938

939

940

941

942

943

944

945

946

947

948

949

950

951

952

953

954

955

956

957

958

959

960

961

962

963

964

965

966

967

968

969

970

971

972

973

974

975

976

977

978

979

980

981

982

983

984

985

986

987

988

989

990

991

992

993

994

995

996

997

998

999

1000

XX New isolated LAGE-1 tumour associated nucleic acids - used to
 PT develop products for the diagnosis and treatment of LAGE-1
 PT associated disorders, particularly tumours
 XX
 PS claim 1: Page 50 52, 73pp; English.
 XX
 CC The present sequence encodes LAGE-1 tumour associated protein (TAP).
 CC The present invention also describes: (1) a method for treating a
 CC subject with a disorder characterised by expression of a LAGE-1 nucleic
 CC acid molecule or an expression product, comprising administering to the
 CC subject autologous cytolytic T cells to ameliorate the disorder, where
 CC the cytolytic T cells are specific for complexes of an HLA molecule and
 CC a LAGE-1 TAP or an immunogenic fragment, (2) a method for treating a
 CC subject with a disorder characterised by expression of a LAGE-1 nucleic
 CC acid molecule or an expression product, comprising administering a
 CC LAGE-1 TAP or an immunogenic fragment to ameliorate the disorder; and
 CC (3) a method for selectively enriching a population of T cells with
 CC cytolytic T cells specific for a LAGE-1 TAP comprising contacting an
 CC isolated population of T cells with an agent presenting a complex of a
 CC LAGE TAP or an immunogenic fragment and a HLA presenting molecule to
 CC selectively enrich the isolated population of T cells with the cytolytic
 CC T cells. The methods and products from the present invention can be used
 CC for the diagnosis and treatment of LAGE-1 associated disorders,
 CC particularly tumours.
 XX
 SQ Sequence 1002 BP: 170 A: 302 C: 336 G: 185 T: 6 other:

Query Match 100 0%; Score 1002; DB 19; Length 1002;
 Best Local Similarity 100 0%; Prod. No. 1a 215;
 Matches 1002; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 tctgctccatctctgtggccctgacattctctctgagagcggcgagaggtccgg 60
 Db 1 tctgctccatctctgtggccctgacattctctctgagagcggcgagaggtccgg 60
 QY 61 atccatgagagcagagcagagcagagcagagcagagcagagcagagcagagc 120
 Db 61 atccatgagagcagagcagagcagagcagagcagagcagagcagagcagagc 120
 QY 121 agacccagcattctctgagcagagcagagcagagcagagcagagcagagc 180
 Db 121 agacccagcattctctgagcagagcagagcagagcagagcagagcagagc 180
 QY 181 cagagcagagcagagcagagcagagcagagcagagcagagcagagcagagc 240
 Db 181 cagagcagagcagagcagagcagagcagagcagagcagagcagagcagagc 240
 QY 241 cagagcagagcagagcagagcagagcagagcagagcagagcagagcagagc 300
 Db 241 cagagcagagcagagcagagcagagcagagcagagcagagcagagcagagc 300
 QY 301 ggcagagcagagcagagcagagcagagcagagcagagcagagcagagc 360
 Db 301 ggcagagcagagcagagcagagcagagcagagcagagcagagcagagc 360
 QY 361 catcaaaacagagcagagcagagcagagcagagcagagcagagcagagc 420
 Db 361 catcaaaacagagcagagcagagcagagcagagcagagcagagcagagc 420
 QY 421 agagcagagcagagcagagcagagcagagcagagcagagcagagcagagc 480
 Db 421 agagcagagcagagcagagcagagcagagcagagcagagcagagcagagc 480
 QY 481 cagagcagagcagagcagagcagagcagagcagagcagagcagagcagagc 540
 Db 481 cagagcagagcagagcagagcagagcagagcagagcagagcagagcagagc 540
 QY 541 cagagcagagcagagcagagcagagcagagcagagcagagcagagcagagc 600
 Db 541 cagagcagagcagagcagagcagagcagagcagagcagagcagagcagagc 600

QY 601 accatgagagcagagcagagcagagcagagcagagcagagcagagcagagc 660
 Db 601 accatgagagcagagcagagcagagcagagcagagcagagcagagcagagc 660
 QY 661 cgcctttaaataatgattctctgacccctacatttaccctgactgactgacagacc 720
 Db 661 cgcctttaaataatgattctctgacccctacatttaccctgactgactgacagacc 720
 QY 721 gccactgagagcagagcagagcagagcagagcagagcagagcagagcagagc 780
 Db 721 gccactgagagcagagcagagcagagcagagcagagcagagcagagcagagc 780
 QY 781 cgcagatcttctgcccgtgatttttgggtcaggtccctcagagcagagcagagc 840
 Db 781 cgcagatcttctgcccgtgatttttgggtcaggtccctcagagcagagcagagc 840
 QY 841 adccctcagagcagagcagagcagagcagagcagagcagagcagagcagagc 900
 Db 841 adccctcagagcagagcagagcagagcagagcagagcagagcagagcagagc 900
 QY 901 gccagttcatttggggggtgatttttgggtcaggtccctcagagcagagcagagc 960
 Db 901 gccagttcatttggggggtgatttttgggtcaggtccctcagagcagagcagagc 960
 QY 961 ttctgagagaaataaagctgagctacgtattcccaaaaaa 1002
 Db 961 ttctgagagaaataaagctgagctacgtattcccaaaaaa 1002

RESULT 2
 ID D00151 standard; cDNA: 993 BP.
 XX AC D00151;
 XX DT 31-JUL-2000 (first entry)
 XX DE Human LAGE-1 unsplined variant, LAGE-11 encoding cDNA.
 XX KW LAGE-11; CAMEL, CTL-recognised Antigen on MELANOMA; human; cancer; CTL;
 KW cytotoxic T lymphocyte; tumour-associated antigen; NY-ESO-1; anticancer;
 KW melanoma; immunotherapy; immune response; ss.
 XX OS Homo sapiens.
 XX FH Key Location/Qualifiers
 FT 5'UTR 1..55
 FT CDS /*tag= a
 FT /*tag= b
 FT /*tag= c
 FT /*tag= d
 FT /*tag= e
 FT /*tag= f
 FT /*tag= g
 FT /*tag= h
 FT /*tag= i
 FT /*tag= j
 FT /*tag= k
 FT /*tag= l
 FT /*tag= m
 FT /*tag= n
 FT /*tag= o
 FT /*tag= p
 FT /*tag= q
 FT /*tag= r
 FT /*tag= s
 FT /*tag= t
 FT /*tag= u
 FT /*tag= v
 FT /*tag= w
 FT /*tag= x
 FT /*tag= y
 FT /*tag= z
 FT /*tag= AA
 FT /*tag= AB
 FT /*tag= AC
 FT /*tag= AD
 FT /*tag= AE
 FT /*tag= AF
 FT /*tag= AG
 FT /*tag= AH
 FT /*tag= AI
 FT /*tag= AJ
 FT /*tag= AK
 FT /*tag= AL
 FT /*tag= AM
 FT /*tag= AN
 FT /*tag= AO
 FT /*tag= AP
 FT /*tag= AQ
 FT /*tag= AR
 FT /*tag= AS
 FT /*tag= AT
 FT /*tag= AU
 FT /*tag= AV
 FT /*tag= AW
 FT /*tag= AX
 FT /*tag= AY
 FT /*tag= AZ
 FT /*tag= BA
 FT /*tag= BB
 FT /*tag= BC
 FT /*tag= BD
 FT /*tag= BE
 FT /*tag= BF
 FT /*tag= BG
 FT /*tag= BH
 FT /*tag= BI
 FT /*tag= BJ
 FT /*tag= BK
 FT /*tag= BL
 FT /*tag= BM
 FT /*tag= BN
 FT /*tag= BO
 FT /*tag= BP
 FT /*tag= BQ
 FT /*tag= BR
 FT /*tag= BS
 FT /*tag= BT
 FT /*tag= BU
 FT /*tag= BV
 FT /*tag= BW
 FT /*tag= BX
 FT /*tag= BY
 FT /*tag= BZ
 FT /*tag= CA
 FT /*tag= CB
 FT /*tag= CC
 FT /*tag= CD
 FT /*tag= CE
 FT /*tag= CF
 FT /*tag= CG
 FT /*tag= CH
 FT /*tag= CI
 FT /*tag= CJ
 FT /*tag= CK
 FT /*tag= CL
 FT /*tag= CM
 FT /*tag= CN
 FT /*tag= CO
 FT /*tag= CP
 FT /*tag= CQ
 FT /*tag= CR
 FT /*tag= CS
 FT /*tag= CT
 FT /*tag= CU
 FT /*tag= CV
 FT /*tag= CW
 FT /*tag= CX
 FT /*tag= CY
 FT /*tag= CZ
 FT /*tag= DA
 FT /*tag= DB
 FT /*tag= DC
 FT /*tag= DD
 FT /*tag= DE
 FT /*tag= DF
 FT /*tag= DG
 FT /*tag= DH
 FT /*tag= DI
 FT /*tag= DJ
 FT /*tag= DK
 FT /*tag= DL
 FT /*tag= DM
 FT /*tag= DN
 FT /*tag= DO
 FT /*tag= DP
 FT /*tag= DQ
 FT /*tag= DR
 FT /*tag= DS
 FT /*tag= DT
 FT /*tag= DU
 FT /*tag= DV
 FT /*tag= DW
 FT /*tag= DX
 FT /*tag= DY
 FT /*tag= DZ
 FT /*tag= EA
 FT /*tag= EB
 FT /*tag= EC
 FT /*tag= ED
 FT /*tag= EE
 FT /*tag= EF
 FT /*tag= EG
 FT /*tag= EH
 FT /*tag= EI
 FT /*tag= EJ
 FT /*tag= EK
 FT /*tag= EL
 FT /*tag= EM
 FT /*tag= EN
 FT /*tag= EO
 FT /*tag= EP
 FT /*tag= EQ
 FT /*tag= ER
 FT /*tag= ES
 FT /*tag= ET
 FT /*tag= EU
 FT /*tag= EV
 FT /*tag= EW
 FT /*tag= EX
 FT /*tag= EY
 FT /*tag= EZ
 FT /*tag= FA
 FT /*tag= FB
 FT /*tag= FC
 FT /*tag= FD
 FT /*tag= FE
 FT /*tag= FF
 FT /*tag= FG
 FT /*tag= FH
 FT /*tag= FI
 FT /*tag= FJ
 FT /*tag= FK
 FT /*tag= FL
 FT /*tag= FM
 FT /*tag= FN
 FT /*tag= FO
 FT /*tag= FP
 FT /*tag= FQ
 FT /*tag= FR
 FT /*tag= FS
 FT /*tag= FT
 FT /*tag= FU
 FT /*tag= FV
 FT /*tag= FW
 FT /*tag= FX
 FT /*tag= FY
 FT /*tag= FZ
 FT /*tag= GA
 FT /*tag= GB
 FT /*tag= GC
 FT /*tag= GD
 FT /*tag= GE
 FT /*tag= GF
 FT /*tag= GG
 FT /*tag= GH
 FT /*tag= GI
 FT /*tag= GJ
 FT /*tag= GK
 FT /*tag= GL
 FT /*tag= GM
 FT /*tag= GN
 FT /*tag= GO
 FT /*tag= GP
 FT /*tag= GQ
 FT /*tag= GR
 FT /*tag= GS
 FT /*tag= GT
 FT /*tag= GU
 FT /*tag= GV
 FT /*tag= GW
 FT /*tag= GX
 FT /*tag= GY
 FT /*tag= GZ
 FT /*tag= HA
 FT /*tag= HB
 FT /*tag= HC
 FT /*tag= HD
 FT /*tag= HE
 FT /*tag= HF
 FT /*tag= HG
 FT /*tag= HH
 FT /*tag= HI
 FT /*tag= HJ
 FT /*tag= HK
 FT /*tag= HL
 FT /*tag= HM
 FT /*tag= HN
 FT /*tag= HO
 FT /*tag= HP
 FT /*tag= HQ
 FT /*tag= HR
 FT /*tag= HS
 FT /*tag= HT
 FT /*tag= HU
 FT /*tag= HV
 FT /*tag= HW
 FT /*tag= HX
 FT /*tag= HY
 FT /*tag= HZ
 FT /*tag= IA
 FT /*tag= IB
 FT /*tag= IC
 FT /*tag= ID
 FT /*tag= IE
 FT /*tag= IF
 FT /*tag= IG
 FT /*tag= IH
 FT /*tag= II
 FT /*tag= IJ
 FT /*tag= IK
 FT /*tag= IL
 FT /*tag= IM
 FT /*tag= IN
 FT /*tag= IO
 FT /*tag= IP
 FT /*tag= IQ
 FT /*tag= IR
 FT /*tag= IS
 FT /*tag= IT
 FT /*tag= IU
 FT /*tag= IV
 FT /*tag= IW
 FT /*tag= IX
 FT /*tag= IY
 FT /*tag= IZ
 FT /*tag= JA
 FT /*tag= JB
 FT /*tag= JC
 FT /*tag= JD
 FT /*tag= JE
 FT /*tag= JF
 FT /*tag= JG
 FT /*tag= JH
 FT /*tag= JI
 FT /*tag= JJ
 FT /*tag= JK
 FT /*tag= JL
 FT /*tag= JM
 FT /*tag= JN
 FT /*tag= JO
 FT /*tag= JP
 FT /*tag= JQ
 FT /*tag= JR
 FT /*tag= JS
 FT /*tag= JT
 FT /*tag= JU
 FT /*tag= JV
 FT /*tag= JW
 FT /*tag= JX
 FT /*tag= JY
 FT /*tag= JZ
 FT /*tag= KA
 FT /*tag= KB
 FT /*tag= KC
 FT /*tag= KD
 FT /*tag= KE
 FT /*tag= KF
 FT /*tag= KG
 FT /*tag= KH
 FT /*tag= KI
 FT /*tag= KJ
 FT /*tag= KK
 FT /*tag= KL
 FT /*tag= KM
 FT /*tag= KN
 FT /*tag= KO
 FT /*tag= KP
 FT /*tag= KQ
 FT /*tag= KR
 FT /*tag= KS
 FT /*tag= KT
 FT /*tag= KU
 FT /*tag= KV
 FT /*tag= KW
 FT /*tag= KX
 FT /*tag= KY
 FT /*tag= KZ
 FT /*tag= LA
 FT /*tag= LB
 FT /*tag= LC
 FT /*tag= LD
 FT /*tag= LE
 FT /*tag= LF
 FT /*tag= LG
 FT /*tag= LH
 FT /*tag= LI
 FT /*tag= LJ
 FT /*tag= LK
 FT /*tag= LL
 FT /*tag= LM
 FT /*tag= LN
 FT /*tag= LO
 FT /*tag= LP
 FT /*tag= LQ
 FT /*tag= LR
 FT /*tag= LS
 FT /*tag= LT
 FT /*tag= LU
 FT /*tag= LV
 FT /*tag= LW
 FT /*tag= LX
 FT /*tag= LY
 FT /*tag= LZ
 FT /*tag= MA
 FT /*tag= MB
 FT /*tag= MC
 FT /*tag= MD
 FT /*tag= ME
 FT /*tag= MF
 FT /*tag= MG
 FT /*tag= MH
 FT /*tag= MI
 FT /*tag= MJ
 FT /*tag= MK
 FT /*tag= ML
 FT /*tag= MM
 FT /*tag= MN
 FT /*tag= MO
 FT /*tag= MP
 FT /*tag= MQ
 FT /*tag= MR
 FT /*tag= MS
 FT /*tag= MT
 FT /*tag= MU
 FT /*tag= MV
 FT /*tag= MW
 FT /*tag= MX
 FT /*tag= MY
 FT /*tag= MZ
 FT /*tag= NA
 FT /*tag= NB
 FT /*tag= NC
 FT /*tag= ND
 FT /*tag= NE
 FT /*tag= NF
 FT /*tag= NG
 FT /*tag= NH
 FT /*tag= NI
 FT /*tag= NJ
 FT /*tag= NK
 FT /*tag= NL
 FT /*tag= NM
 FT /*tag= NN
 FT /*tag= NO
 FT /*tag= NP
 FT /*tag= NQ
 FT /*tag= NR
 FT /*tag= NS
 FT /*tag= NT
 FT /*tag= NU
 FT /*tag= NV
 FT /*tag= NW
 FT /*tag= NX
 FT /*tag= NY
 FT /*tag= NZ
 FT /*tag= OA
 FT /*tag= OB
 FT /*tag= OC
 FT /*tag= OD
 FT /*tag= OE
 FT /*tag= OF
 FT /*tag= OG
 FT /*tag= OH
 FT /*tag= OI
 FT /*tag= OJ
 FT /*tag= OK
 FT /*tag= OL
 FT /*tag= OM
 FT /*tag= ON
 FT /*tag= OO
 FT /*tag= OP
 FT /*tag= OQ
 FT /*tag= OR
 FT /*tag= OS
 FT /*tag= OT
 FT /*tag= OU
 FT /*tag= OV
 FT /*tag= OW
 FT /*tag= OX
 FT /*tag= OY
 FT /*tag= OZ
 FT /*tag= PA
 FT /*tag= PB
 FT /*tag= PC
 FT /*tag= PD
 FT /*tag= PE
 FT /*tag= PF
 FT /*tag= PG
 FT /*tag= PH
 FT /*tag= PI
 FT /*tag= PJ
 FT /*tag= PK
 FT /*tag= PL
 FT /*tag= PM
 FT /*tag= PN
 FT /*tag= PO
 FT /*tag= PP
 FT /*tag= PQ
 FT /*tag= PR
 FT /*tag= PS
 FT /*tag= PT
 FT /*tag= PU
 FT /*tag= PV
 FT /*tag= PW
 FT /*tag= PX
 FT /*tag= PY
 FT /*tag= PZ
 FT /*tag= QA
 FT /*tag= QB
 FT /*tag= QC
 FT /*tag= QD
 FT /*tag= QE
 FT /*tag= QF
 FT /*tag= QG
 FT /*tag= QH
 FT /*tag= QI
 FT /*tag= QJ
 FT /*tag= QK
 FT /*tag= QL
 FT /*tag= QM
 FT /*tag= QN
 FT /*tag= QO
 FT /*tag= QP
 FT /*tag= QQ
 FT /*tag= QR
 FT /*tag= QS
 FT /*tag= QT
 FT /*tag= QU
 FT /*tag= QV
 FT /*tag= QW
 FT /*tag= QX
 FT /*tag= QY
 FT /*tag= QZ
 FT /*tag= RA
 FT /*tag= RB
 FT /*tag= RC
 FT /*tag= RD
 FT /*tag= RE
 FT /*tag= RF
 FT /*tag= RG
 FT /*tag= RH
 FT /*tag= RI
 FT /*tag= RJ
 FT /*tag= RK
 FT /*tag= RL
 FT /*tag= RM
 FT /*tag= RN
 FT /*tag= RO
 FT /*tag= RP
 FT /*tag= RQ
 FT /*tag= RR
 FT /*tag= RS
 FT /*tag= RT
 FT /*tag= RU
 FT /*tag= RV
 FT /*tag= RW
 FT /*tag= RX
 FT /*tag= RY
 FT /*tag= RZ
 FT /*tag= SA
 FT /*tag= SB
 FT /*tag= SC
 FT /*tag= SD
 FT /*tag= SE
 FT /*tag= SF
 FT /*tag= SG
 FT /*tag= SH
 FT /*tag= SI
 FT /*tag= SJ
 FT /*tag= SK
 FT /*tag= SL
 FT /*tag= SM
 FT /*tag= SN
 FT /*tag= SO
 FT /*tag= SP
 FT /*tag= SQ
 FT /*tag= SR
 FT /*tag= SS
 FT /*tag= ST
 FT /*tag= SU
 FT /*tag= SV
 FT /*tag= SW
 FT /*tag= SX
 FT /*tag= SY
 FT /*tag= SZ
 FT /*tag= TA
 FT /*tag= TB
 FT /*tag= TC
 FT /*tag= TD
 FT /*tag= TE
 FT /*tag= TF
 FT /*tag= TG
 FT /*tag= TH
 FT /*tag= TI
 FT /*tag= TJ
 FT /*tag= TK
 FT /*tag= TL
 FT /*tag= TM
 FT /*tag= TN
 FT /*tag= TO
 FT /*tag= TP
 FT /*tag= TQ
 FT /*tag= TR
 FT /*tag= TS
 FT /*tag= TU
 FT /*tag= TV
 FT /*tag= TW
 FT /*tag= TX
 FT /*tag= TY
 FT /*tag= TZ
 FT /*tag= UA
 FT /*tag= UB
 FT /*tag= UC
 FT /*tag= UD
 FT /*tag= UE
 FT /*tag= UF
 FT /*tag= UG
 FT /*tag= UH
 FT /*tag= UI
 FT /*tag= UJ
 FT /*tag= UK
 FT /*tag= UL
 FT /*tag= UM
 FT /*tag= UN
 FT /*tag= UO
 FT /*tag= UP
 FT /*tag= UQ
 FT /*tag= UR
 FT /*tag= US
 FT /*tag= UT
 FT /*tag= UV
 FT /*tag= UW
 FT /*tag= UX
 FT /*tag= UY
 FT /*tag= UZ
 FT /*tag= VA
 FT /*tag= VB
 FT /*tag= VC
 FT /*tag= VD
 FT /*tag= VE
 FT /*tag= VF
 FT /*tag= VG
 FT /*tag= VH
 FT /*tag= VI
 FT /*tag= VJ
 FT /*tag= VK
 FT /*tag= VL
 FT /*tag= VM
 FT /*tag= VN
 FT /*tag= VO
 FT /*tag= VP
 FT /*tag= VQ
 FT /*tag= VR
 FT /*tag= VS
 FT /*tag= VT
 FT /*tag= VU
 FT /*tag= VV
 FT /*tag= VW
 FT /*tag= VX
 FT /*tag= VY
 FT /*tag= VZ
 FT /*tag= WA
 FT /*tag= WB
 FT /*tag= WC
 FT /*tag= WD
 FT /*tag= WE
 FT /*tag= WF
 FT /*tag= WG
 FT /*tag= WH
 FT /*tag= WI
 FT /*tag= WJ
 FT /*tag= WK
 FT /*tag= WL
 FT /*tag= WM
 FT /*tag= WN
 FT /*tag= WO
 FT /*tag= WP
 FT /*tag= WQ
 FT /*tag= WR
 FT /*tag= WS
 FT /*tag= WT
 FT /*tag= WU
 FT /*tag= WV
 FT /*tag= WW
 FT /*tag= WX
 FT /*tag= WY
 FT /*tag= WZ
 FT /*tag= XA
 FT /*tag= XB
 FT /*tag= XC
 FT /*tag= XD
 FT /*tag= XE
 FT /*tag= XF
 FT /*tag= XG
 FT /*tag= XH
 FT /*tag= XI
 FT /*tag= XJ
 FT /*tag= XK
 FT /*tag= XL
 FT /*tag= XM
 FT /*tag= XN
 FT /*tag= XO
 FT /*tag= XP
 FT /*tag= XQ
 FT /*tag= XR
 FT /*tag= XS
 FT /*tag= XT
 FT /*tag= XU
 FT /*tag= XV
 FT /*tag= XW
 FT /*tag= XX
 FT /*tag= XY
 FT /*tag= XZ
 FT /*tag= YA
 FT /*tag= YB
 FT /*tag= YC
 FT /*tag= YD
 FT /*tag= YE
 FT /*tag= YF
 FT /*tag= YG
 FT /*tag= YH
 FT /*tag= YI
 FT /*tag= YJ
 FT /*tag= YK
 FT /*tag= YL
 FT /*tag= YM
 FT /*tag= YN
 FT /*tag= YO
 FT /*tag= YP
 FT /*tag= YQ
 FT /*tag= YR
 FT /*tag= YS
 FT /*tag= YT
 FT /*tag= YU
 FT /*tag= YV
 FT /*tag= YW
 FT /*tag= YX
 FT /*tag= YY
 FT /*tag= YZ
 FT /*tag= ZA
 FT /*tag= ZB
 FT /*tag= ZC
 FT /*tag= ZD
 FT /*tag= ZE
 FT /*tag= ZF
 FT /*tag= ZG
 FT /*tag= ZH
 FT /*tag= ZI
 FT /*tag= ZJ
 FT /*tag= ZK
 FT /*tag= ZL
 FT /*tag= ZM
 FT /*tag= ZN
 FT /*tag= ZO
 FT /*tag= ZP
 FT /*tag= ZQ
 FT /*tag= ZR
 FT /*tag= ZS
 FT /*tag= ZT
 FT /*tag= ZU
 FT /*tag= ZV
 FT /*tag= ZW
 FT /*tag= ZX
 FT /*tag= ZY
 FT /*tag= ZZ

WT1: 2000-339685/29.
 DR: 3: F3CB, Y70861.
 PT Tumor-associated antigen useful for cancer immunotherapy is encoded by the open reading frame of LAGE-1 (a tumor-specific antigen) cDNA -


```

Db 121 ttcctgatgcacagggagcattatctgctggccagatagagggctgtccacaggggca 180
QY 192 gaggcccccagggagagagagagagagagagagagagagagagagagagagagagag 251
Db 181 gaaatcccccagggagagagagagagagagagagagagagagagagagagagagag 246
QY 252 atcccatcagagagagagagagagagagagagagagagagagagagagagagagag 311
Db 243 atcccatcagagagagagagagagagagagagagagagagagagagagagagagag 304
QY 312 cagagagagagagagagagagagagagagagagagagagagagagagagagagagag 371
Db 301 cagagagagagagagagagagagagagagagagagagagagagagagagagagagag 360
QY 372 agctatcagagagagagagagagagagagagagagagagagagagagagagagag 431
Db 361 agctatcagagagagagagagagagagagagagagagagagagagagagagagag 420
QY 432 atagagagagagagagagagagagagagagagagagagagagagagagagagagag 491
Db 421 atagagagagagagagagagagagagagagagagagagagagagagagagagagag 457
QY 492 atagagagagagagagagagagagagagagagagagagagagagagagagagagag 551
Db 458 ----- 457
QY 552 atagagagagagagagagagagagagagagagagagagagagagagagagagagag 611
Db 458 ----- 457
QY 612 cagagagagagagagagagagagagagagagagagagagagagagagagagagag 671
Db 458 ----- 457
QY 672 atatttctctgcctccatctttagccagagagagagagagagagagagagagagag 731
Db 458 ----- 457
QY 732 atatttctctgcctccatctttagccagagagagagagagagagagagagagagag 791
Db 492 ctctccatcagctctctctccagagagagagagagagagagagagagagagagagag 551
QY 792 ctctccatcagctctctctccagagagagagagagagagagagagagagagagagag 851
Db 552 ctctccatcagctctctctccagagagagagagagagagagagagagagagagagag 611
QY 852 ctctccatcagctctctctccagagagagagagagagagagagagagagagagagag 911
Db 612 ctctccatcagctctctctccagagagagagagagagagagagagagagagagagag 671
QY 912 atagagagagagagagagagagagagagagagagagagagagagagagagagagag 971
Db 672 atagagagagagagagagagagagagagagagagagagagagagagagagagagag 731
QY 972 atagagagagagagagagagagagagagagagagagagagagagagagagagagag 988
Db 732 atagagagagagagagagagagagagagagagagagagagagagagagagagagag 748

```

RESULT 8

D00152
ID D00152 standard: cDNA; 752 BP.

XX AC D00152:

XX D00152:

DI 31-JUL-2000 (first entry)

DE Human tumour antigen, NY-ESO-1 cDNA.

KW NY ESO-1; CAMEL, CTL recognised Antigen on MELANOMA; Human; cancer; CTL;
KW cytotoxic T lymphocyte, tumour-associated antigen; LAGE-1, anti-tumour,
KW melanoma, immunotherapy, immune response; ss.

OS Homo sapiens.

XX Key Location/Qualifiers
FT 5'UTR 1..53 /tag= a
FT CDS 54..596 /tag= b
FT /product= "Human NY-ESO-1 protein"
FT /note= "derived from open reading frame (ORF)-1"
FT 3'UTR 597..752 /tag= c
FT 5'UTR 1..93 /tag= d
FT /note= "5' UTR of ORF-2"
FT CDS 94..270 /tag= e
FT /product= "Human NY ESO-1 short variant protein"
FT /note= "derived from open reading frame (ORF)-2"
FT 3'UTR 271..752 /tag= f
FT /note= "3' UTR of ORF-2"

W0200002584-A1.

XX 27-APR-2000.

XX 15-OCT-1999; 95WC-LF07832.

XX 16-OCT-1998; 98EP-0119583.

XX (BOEH) BOEHRINGER INGELHEIM INT GMBH.
XX (UYHO-) UNIV HOSPITAL LEIDEN.

XX Schrier PI, Aarnoudse CA, Heider K, Klade C;

XX WP1; 2000-339685/29.

XX E-PSUB; Y/0862, Y/0863.

XX Tumor associated antigen useful for cancer immunotherapy, is encoded by
FT the open reading frame of LAGE-1 (a tumor-specific antigen) cDNA -

XX Example 1; Page 61-62; 73pp; English.

XX The present sequence is the cDNA encoding human NY-ESO-1, a tumour
CC antigen, identified by screening a cDNA library of an esophagus
CC carcinoma. It has two different reading frames and the proteins encoded
CC contain epitopes of tumour specific proteins. NY-ESO-1 is expressed in
CC different tumour types, but not in healthy tissues except in testis. It
CC also shows homology with the CAMEL (cytotoxic T lymphocytes (CTL)-
CC recognised Antigen on MELANOMA) cDNA, a tumour-associated antigen. The
CC tumour-associated antigen displayed on melanoma cells is recognised by
CC cytotoxic T lymphocytes. This sequence has anticancer activity. CAMEL
CC tumour antigen and immunogenic peptides derived from it are useful for
CC cancer immunotherapy. They have the potential to induce an immune
CC response, by eliciting a CTL response. The DNA molecule is used for the
CC construction of recombinant or fusion proteins

XX Sequence 752 BP, 126 A, 236 C, 256 G, 149 T; 0 other;

Query Match 43.8%; Score 438.6; DB 21; Length 752;
Best Local Similarity 72.1%; Pred. No. 1.5e-89;

Matches 704; Conservative 0; Mismatches 44; Indels 239; Gaps 1;

QY 12 atctctgag 71

Db 1 atctctgag 60

QY 72 cag 131

Db 61 cag 120

QY 132 ttcctgatgcacagggagcattatctgctggccagatagagggctgtccacaggggca 191

Gapcore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

em nucleide nucleide search using sw model

Run on: May 1, 2001, 09:37:47 ; Search time 2512.88 Seconds

(without alignments)

5880.645 Million cell updates/sec

Hit for: US-09-341-829A.4

Posttest scores: 1002

Sequences: 1 tetapctccatctctcgtq.....ctacqatccqaaaaaaa 1002

Scoring table: IDENTIFY_WEP

Gapop 10.0 ; Gapext 1.0

Searches: 1284245 seqs, 7474929652 residues

Total number of hits satisfying chosen parameters: 2566470

Minimum DB seq length: 6

Maximum DB seq length: 2000000000

Post processing: Minimum Match 0%

Maximum Match 100%

Listing first 47 summaries

Database :

Genbank :

1: gb_ba1.*

2: gb_ba2.*

3: gb_ba3.*

4: gb_ba4.*

5: gb_ba5.*

6: gb_ba6.*

7: gb_ba7.*

8: gb_ba8.*

9: gb_ba9.*

10: gb_ba10.*

11: gb_ba11.*

12: gb_ba12.*

13: gb_ba13.*

14: gb_ba14.*

15: gb_ba15.*

16: gb_ba16.*

17: gb_ba17.*

18: gb_ba18.*

19: gb_ba19.*

20: gb_ba20.*

21: gb_ba21.*

22: gb_ba22.*

23: gb_ba23.*

24: gb_ba24.*

25: gb_ba25.*

26: gb_ba26.*

27: gb_ba27.*

28: gb_ba28.*

29: gb_ba29.*

30: gb_ba30.*

31: gb_ba31.*

32: gb_ba32.*

33: gb_ba33.*

34: gb_ba34.*

35: gb_ba35.*

36: gb_ba36.*

37: gb_ba37.*

38: gb_ba38.*

39: gb_ba39.*

40: gb_ba40.*

41: gb_ba41.*

42: gb_ba42.*

43: gb_ba43.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1002	100.0	1002	9	AK042545
2	1002	100.0	1002	92	AK042545
3	985	98.3	993	9	AX024689
4	985	98.3	993	92	AX024689
5	706.8	70.5	873	92	HS012835
6	66.0	65.9	112874	78	HS012835
7	656.8	65.5	3245	91	HS012835
8	540	53.9	2640	92	HS012835
9	540	53.9	112874	78	HS012835
10	515	51.4	767	9	AX024687
11	515	51.4	767	92	AX024687

ORIGIN

Query Match: 98.38; Score 985; DB 9; Length 993;
Best Local Similarity 99.58; Prod. No. 1.2e-170;
Matches 988; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Q7 10 acatctctatgaagctgaactctctctctgaagcaggcagaagctccgaagccatata 69
DB 1 GUAICTGCGGCGCTGCACTCTCTCTGACAGCTGGGAGAGTGGGAGCCATGGA 69
Q7 70 gggag 129
DB 61 GGGCAAG 120
Q7 130 catctctctatgaagcagaagcagaagcagaagcagaagcagaagcagaagcaga 189
DB 121 TATCTCTATGACGAGAGGCAATGCTGGCGGAGAGAGAGAGAGAGAGAGAGAG 180
Q7 150 caaagctcccccaggagagagagagagagagagagagagagagagagagagag 249
DB 181 TAAAGGTCGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
Q7 250 aaatccagatgaagctgaagctctctgaagcagaagcagaagcagaagcagaag 309
DB 241 AGGTCGAGATGCGAG 300
Q7 310 gggag 369
DB 401 GAG 360
Q7 370 gggag 429
DB 461 GAG 420
Q7 440 gggag 459
DB 421 TCGAAGAGATTCACCGTGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
Q7 490 gggag 549
DB 481 GAG 540
Q7 550 gggag 609
DB 541 GAG 600
Q7 610 gggag 669
DB 601 AGGAG 660
Q7 670 gggag 729
DB 661 TGTATGTTCTGCGGCTGACATTTAGCGGAGTGGTGGAGAGAGAGAGAGAGAGAG 720
Q7 740 gggag 789
DB 741 AGCTGCTGAG 780
Q7 790 gggag 849
DB 781 TGTGCGAG 840
Q7 850 gggag 909
DB 841 GAG 900
Q7 910 gggag 969
DB 901 TGTGCGAG 960
Q7 976 gggag 1002
|||||

Db 951 AATTAAGCTCAGCTACGAAAAAAGAAAAA 993

RESULT 4

HSA012835

LOCUS

DEFINITION

Accession

Version

Keywords

Source

Organism

Reference

Authors

Title

Journal

Milline

Reference

Submitted

University

Netherlands

Location/Qualifiers

Source

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

Query Match: 98.38; Score 985; DB 92; Length 993;
Best Local Similarity 99.58; Prod. No. 1.2e-170;
Matches 988; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 10 gcatctctatgaagctgaactctctctctgaagcaggcagaagctccgaagccatata 993

Genome version 4.5
Copyright (c) 1994 - 2000 CompuGen Ltd.

CM nucleotide nucleotide search, using sw model

Run on: April 05, 2001, 24:54:02 : Search time 4786.75 Seconds

(without alignments)

2311631 Billion cell updates/sec

File: us-09-341-829a-4

Port: 2002

Sequence: 1 tctgctccgacatctctctg.....ctacgattccgagagagaa 1002

Scoring table: HUPNUTLY NUC

Gapop 10.00, Gapext 1.0

Searched: 9625377 seqs, 436804070 residues

Total number of hits satisfying chosen parameters: 19247034

Minimum hit seq length: 0

Maximum hit seq length: 24000000000

Post processing: Maximum Match 0%

Listing first 45 summaries

Database:

EST:

1: qb_est1: *
2: qb_est2: *
3: qb_est3: *
4: qb_est4: *
5: qb_est5: *
6: qb_est6: *
7: qb_est7: *
8: qb_est8: *
9: qb_est9: *
10: qb_est10: *
11: qb_est11: *
12: qb_est12: *
13: qb_est13: *
14: qb_est14: *
15: qb_est15: *
16: qb_est16: *
17: qb_est17: *
18: qb_est18: *
19: qb_est19: *
20: qb_est20: *
21: qb_est21: *
22: qb_est22: *
23: qb_est23: *
24: qb_est24: *
25: qb_est25: *
26: qb_est26: *
27: qb_est27: *
28: qb_est28: *
29: qb_est29: *
30: qb_est30: *
31: qb_est31: *
32: qb_est32: *
33: qb_est33: *
34: qb_est34: *
35: qb_est35: *
36: qb_est36: *
37: qb_est37: *
38: qb_est38: *
39: qb_est39: *
40: qb_est40: *
41: qb_est41: *
42: qb_est42: *
43: qb_est43: *
44: qb_est44: *
45: qb_est45: *

44: em_esthum10: *
45: em_esthum11: *
46: em_esthum12: *
47: em_esthum13: *
48: em_esthum14: *
49: em_esthum15: *
50: em_esthum16: *
51: em_esthum17: *
52: em_esthum18: *
53: em_esthum19: *
54: em_esthum20: *
55: em_esthum21: *
56: em_esthum22: *
57: em_esthum23: *
58: em_esthum24: *
59: em_esthum25: *
60: em_esthum26: *
61: em_esthum27: *
62: em_esthum28: *
63: em_esthum29: *
64: em_esthum30: *
65: em_esthum31: *
66: em_esthum32: *
67: em_esthum33: *
68: em_esthum34: *
69: em_esthum35: *
70: em_esthum36: *
71: em_esthum37: *
72: em_esthum38: *
73: em_esthum39: *
74: em_esthum40: *
75: em_esthum41: *
76: em_esthum42: *
77: em_esthum43: *
78: em_esthum44: *
79: em_esthum45: *
80: em_esthum46: *
81: em_esthum47: *
82: em_esthum48: *
83: em_esthum49: *
84: em_esthum50: *
85: em_esthum51: *
86: em_esthum52: *
87: em_esthum53: *
88: em_esthum54: *
89: em_esthum55: *
90: em_esthum56: *
91: em_esthum57: *
92: em_esthum58: *
93: em_esthum59: *
94: em_esthum60: *
95: em_esthum61: *
96: em_esthum62: *
97: em_esthum63: *
98: em_esthum64: *
99: em_esthum65: *
100: em_esthum66: *
101: em_esthum67: *
102: em_esthum68: *
103: em_esthum69: *
104: em_esthum70: *
105: em_esthum71: *
106: em_esthum72: *
107: em_esthum73: *
108: em_esthum74: *
109: em_esthum75: *
110: em_esthum76: *
111: em_esthum77: *
112: em_esthum78: *
113: em_esthum79: *
114: em_esthum80: *
115: em_esthum81: *
116: em_esthum82: *

```

117: qb_est48:*
118: qb_est49:*
119: qb_est50:*
120: qb_est51:*
121: qb_est52:*
122: qb_est53:*
123: qb_est54:*
124: qb_est55:*
125: qb_est56:*
126: qb_est57:*
127: qb_est58:*
128: qb_est59:*
129: qb_est60:*
130: qb_est61:*
131: qb_est62:*
132: qb_est63:*
133: qb_est64:*
134: qb_est65:*
135: qb_est66:*
136: qb_est67:*
137: qb_est68:*
138: qb_est69:*
139: qb_est70:*
140: qb_est71:*
141: qb_est72:*
142: qb_est73:*
143: qb_est74:*
144: qb_est75:*
145: qb_est76:*
146: qb_est77:*
147: qb_est78:*
148: qb_est79:*
149: qb_est80:*
150: qb_est81:*
151: qb_est82:*
152: qb_est83:*
153: qb_est84:*
154: qb_est85:*
155: qb_est86:*
156: qb_est87:*
157: qb_est88:*
158: qb_est89:*
159: qb_est90:*
160: qb_est91:*
161: qb_est92:*
162: qb_est93:*
163: qb_est94:*
164: qb_est95:*
165: qb_est96:*
166: qb_est97:*
167: qb_est98:*
168: qb_est99:*
169: qb_est100:*
170: qb_est101:*
171: qb_est102:*
172: qb_est103:*
173: qb_est104:*
174: qb_est105:*
175: qb_est106:*
176: qb_est107:*
177: qb_est108:*
178: qb_est109:*
179: qb_est110:*
180: qb_est111:*
181: qb_est112:*
182: qb_est113:*
183: qb_est114:*
184: qb_est115:*
185: qb_est116:*
186: qb_est117:*
187: qb_est118:*
188: qb_est119:*
189: qb_est120:*

```

```

190: em_oss_pln1:*
191: em_oss_pln2:*
192: em_oss_pro:*
193: em_oss_rod1:*
194: em_oss_rod2:*
195: em_oss_rod3:*
196: em_oss_rod4:*
197: em_oss_rod5:*
198: em_oss_vrt1:*
199: em_oss_vrt2:*
200: em_oss_vrt3:*
201: qb_oss1:*
202: qb_oss2:*
203: qb_oss3:*
204: qb_oss4:*
205: qb_oss5:*
206: qb_oss6:*
207: qb_oss7:*
208: qb_oss8:*
209: qb_oss9:*
210: qb_oss10:*
211: qb_oss11:*
212: qb_oss12:*
213: qb_oss13:*
214: qb_oss14:*
215: qb_oss15:*
216: qb_oss16:*
217: qb_oss17:*
218: qb_oss18:*
219: qb_oss19:*
220: qb_oss20:*
221: qb_oss21:*
222: qb_oss22:*
223: qb_oss23:*
224: qb_oss24:*
225: qb_oss25:*
226: qb_oss26:*
227: qb_oss27:*
228: qb_oss28:*
229: qb_oss29:*
230: qb_oss30:*
231: qb_oss31:*
232: qb_oss32:*
233: qb_oss33:*
234: qb_oss34:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result	No.	Score	Query Match	Length	DB	ID	Description
	1	579.2	57.8	629	165	HE274123	HE274123 601120646
	2	528.6	52.8	586	166	HE390259	HE390259 601385141
	3	467.8	46.7	739	166	HE391232	HE391232 601385826
	4	461.8	46.1	559	165	HE276194	HE276194 601144452
	5	393.2	39.2	800	166	HE408892	HE408892 601404769
	6	373.6	37.3	674	166	HE388562	HE388562 601281855
	7	372.2	37.2	646	165	HE410697	HE410697 601400763
	8	357.8	35.7	511	166	HE387918	HE387918 601282166
	9	346.6	34.6	383	166	HE410952	HE410952 601403069
	10	289.2	28.9	359	17	A1218222	A1218222 qb23923.x
	11	269.8	26.9	281	10	AA634317	AA634317 257860518
	12	227.8	22.7	639	166	HE387962	HE387962 601271781
	13	227.6	22.7	952	166	HE385880	HE385880 601275448
	14	86.6	8.6	321	115	AW436516	AW436516 76782 MAR
	15	85.6	8.5	148	13	AA866133	AA866133 qb54309.s
	16	75.8	7.5	429	10	AA682581	AA682581 2120603.s
	17	75.6	7.5	447	154	HE4542	HE4542 2701408.s1
	18	73.2	7.3	693	178	HE727845	HE727845 601564226


```

QY      602  ccttgatgacacagccgcggcgcaccacccaaqaaacaga 641
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
D6      541  CTTGGTACACCAAGCCGCGATCCTCGAGCGAGCATCAGGA 580

RESULT          3
BE391242
FOODS
DEFINITION
    RF391242       739 bp      mRNA           EST
    G01295920F1 RHILM3C_44 Homo sapiens cDNA clone IMAGE:460770H 5'
mRNA sequence.
    BE391242
RF391242.1 GI:9436597
EST.
human.
ORGANISM
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
    1 (bases 1 to 749)
Size: 739 bp (749 nt) lib. J00000
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Tel.: (301) 496-1950
Email: Robert_Strausberg@nih.gov
Tissue Procurement: ATEC
cDNA Library Preparation: Qing Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MSC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:

```

SOURCE: human.
 ORGANISM: Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 749)
 Acc:35211176-acc lib.4007
 REFERENCE: National Institutes of Health, Mammalian Gene Collection (MGC)
 Published (1994)
 Contact: Robert Strausberg, Ph.D.
 Tel: (301) 496-1950
 Email: Robert_Strausberg@nih.gov
 Tissue procurement: ATCC
 cDNA Library Preparation: Ling Hong/Rubin Laboratory
 cDNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LCM259 row: q column: 05
 High quality sequence stop: 631.

```

FEATURES
Source      High quality sequence stop: 631.
Location/Qualifiers
    1..739
        organism="Homo sapiens"
        db_xref="taxon:9606"
        zcclone="IMAGE:3607708"
        zcclone_lib="NH_MGC_44"
        /transcript_coding_start="adrenomedullary cell line"
        /lab_host="DHIOB (phage resistant)"
        note="Organ. Interact Vector; polrta7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GSCA/CAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP CCA Synthesis kit (Stratagene) and SuperScript II RT (Life Technologies)."
BASE COUNT   121 a     213 c     283 g     122 t
ORIGIN
Query Match          46.7%; Score 467.8; DB 166; Length 749;
Host Local Similarity 92.7%; Pred. No. 8,9e-101;
Matches 611; Conservative 0; Mismatches 47; Indels 11; Gaps 11
QY      66  TGGGCGCCTGAACAGCACGTATTCAGCCGCGCATGCTGATGATGATGCCAGGAGGCCC 125
Db      1  TGCAAGCGAACAGCTGAAAGAGGACATAAGAGAGTGACATTGAGTAGAGAGAGG 69
QY      126 CTGGGCTTTCTGATAGCCCCAGGAGGATATCTGTAAGCCAGAGAGAGGATGATGAGCG 185
Db      61  GTGGCATTCCTCATGACTCA-GAGAGAATATGAGAGAGAGGAGAGAGAGAGAGG 119
QY      186 GCCTGAGGAATCCCTCGAGGAGCACAACAGAGAGCTCGAGGCGGAGAGAG-AGGCG 244
Db      120  CGGCGCAGAGGTCGCGGGGCGCAGGGCGCAGGAGAGAGCTGATGAGAGAGAGAG 179
QY      245 CCACAGGAATCCGATAGCATAGCAAGGATCTCTCTCTCTCTCTCTCTCTCTCTCTCT 303
Db      180  CGGCGAGATTCGATAGGCGAGAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 249
QY      304 CAAGAGGCGCGACAGACCGCTGCTCTGATTCAGCATCAGCATGCTTTCTCATGCCCA 362

```



```

10 261 GCAAGTGCAGGTCCTCATCACTGCTGCCAATCAATTTCTTCTGATCTGGATAC 222
10 782 GCAAGTGCAGGTCCTCATCACTGCTGCCAATCAATTTCTTCTGATCTGGATAC 222
10 221 GCAAGTGCAGGTCCTCATCACTGCTGCCAATCAATTTCTTCTGATCTGGATAC 222
10 842 GCAAGTGCAGGTCCTCATCACTGCTGCCAATCAATTTCTTCTGATCTGGATAC 222
10 164 GCAAGTGCAGGTCCTCATCACTGCTGCCAATCAATTTCTTCTGATCTGGATAC 222
10 902 GCAAGTGCAGGTCCTCATCACTGCTGCCAATCAATTTCTTCTGATCTGGATAC 222
10 101 GCAAGTGCAGGTCCTCATCACTGCTGCCAATCAATTTCTTCTGATCTGGATAC 222
10 402 GCAAGTGCAGGTCCTCATCACTGCTGCCAATCAATTTCTTCTGATCTGGATAC 222
10 41 TCTGATCACTGCTGCCAATCAATTTCTTCTGATCTGGATAC 222

```

```

RESULT 12
LOCUS BE387002 639 bp mRNA EST 21-JUL-2000
DEFINITION BE387002.1 NIH_MGC_20 Homo sapiens cDNA clone IMAGE:461545.5.
mRNA sequence.
ACCESSION BE387002
VERSION BE387002.1 GI:9342276
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 639)
AUTHORS NIH-MGC http://www.ncbi.nlm.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Tel: (401) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: ATCC/Wellby/DBP
cDNA Library Preparation: Ling Hong/Robin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: imac.llnl.gov
Plate: LCM280 row: a column: 22
High quality sequence start: 24
High quality sequence stop: 316.
Location/Qualifiers
1..639
/organism="Homo sapiens"
/label="cDNA clone"
/clone_lib="NIH_MGC_20"
/issue_type="mRNA"
/submitter="Robert Strausberg"
/submitter_email="Robert.Strausberg@nih.gov"
/submitter_phone="(401) 496-1550"
/submitter_fax="(401) 496-1550"
/submitter_address="National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Tel: (401) 496-1550
Email: Robert.Strausberg@nih.gov"

```

```

FEATURES
source
1..639
/organism="Homo sapiens"
/label="cDNA clone"
/clone_lib="NIH_MGC_20"
/issue_type="mRNA"
/submitter="Robert Strausberg"
/submitter_email="Robert.Strausberg@nih.gov"
/submitter_phone="(401) 496-1550"
/submitter_fax="(401) 496-1550"
/submitter_address="National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Tel: (401) 496-1550
Email: Robert.Strausberg@nih.gov"

```

```

BASE COUNT 122 a 182 c 233 g 162 t
ORIGIN
Query Match 23.7% Score 237.8; DB 166; Length 639
Best Local Similarity 96.2%; Pred. No. 4, 36-46;
Matches 276; Conservative 0; Mismatches 7; Indels 4 Gaps 4
10 40 GCAAGTGCAGGTCCTCATCACTGCTGCCAATCAATTTCTTCTGATCTGGATAC 222

```

```

10 17 ACAGTGCAGGTCCTCATCACTGCTGCCAATCAATTTCTTCTGATCTGGATAC 222
10 100 GCAAGTGCAGGTCCTCATCACTGCTGCCAATCAATTTCTTCTGATCTGGATAC 222
10 76 GCAAGTGCAGGTCCTCATCACTGCTGCCAATCAATTTCTTCTGATCTGGATAC 222
10 160 GCAAGTGCAGGTCCTCATCACTGCTGCCAATCAATTTCTTCTGATCTGGATAC 222
10 136 GCAAGTGCAGGTCCTCATCACTGCTGCCAATCAATTTCTTCTGATCTGGATAC 222
10 219 GCAAGTGCAGGTCCTCATCACTGCTGCCAATCAATTTCTTCTGATCTGGATAC 222
10 156 GCAAGTGCAGGTCCTCATCACTGCTGCCAATCAATTTCTTCTGATCTGGATAC 222
10 279 GCAAGTGCAGGTCCTCATCACTGCTGCCAATCAATTTCTTCTGATCTGGATAC 222
10 256 GCAAGTGCAGGTCCTCATCACTGCTGCCAATCAATTTCTTCTGATCTGGATAC 222

```

```

BASE COUNT 196 a 250 c 257 g 149 t
ORIGIN
Query Match 23.7% Score 237.8; DB 166; Length 639
Best Local Similarity 96.2%; Pred. No. 4, 36-46;
Matches 276; Conservative 0; Mismatches 7; Indels 4 Gaps 4
10 180 GCAAGTGCAGGTCCTCATCACTGCTGCCAATCAATTTCTTCTGATCTGGATAC 222

```


DB 142 TTTCGGATTAAATTTTGAAGACGCGCCGCAAAAGGGATTTTCAAAAAGCTCTTTTTCG 73
CY 754 agcgaatttccctgtatgatctadgacagcagtgactttctacccgtgcttttttagctcaag 813
DB 11 TT
DB 72 AGGAGCTTTCCTGTTGATGAGGATCCGACAGTCTTTTTCCTCGGTTTTTTTGGCTCAGC 13
CY 814 ctccctcaaggc 825
DB 12 GCGCGGAGGGG 1

Search completed: May 1, 2001, 04:04:56
Job time: 11404 sec.

Patent No. 581519
 GENERAL INFORMATION:
 APPLICANT: Lethy, Bernard
 APPLICANT: Lucus, Sophie
 APPLICANT: de Smet, Charles
 APPLICANT: Godelaine, Danielle
 APPLICANT: Bonn-Falour, Thierry
 TITLE OF INVENTION: Lethy, THUMOR SPECIFIC GENES
 NUMBER OF SEQUENCES: 14
 CORRESPONDENT ADDRESS:
 ADDRESSEE: Well, Greenfield & Sacks, P.C.
 STREET: 600 Atlantic Avenue
 CITY: Boston
 STATE: MA
 COUNTRY: USA
 ZIP: 02210
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 OPERATING SYSTEM: IBM PC compatible
 SOFTWARE: Patent In Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/791,495
 FILING DATE:
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Van Amerongen, John R.
 REGISTRATION NUMBER: 40,212
 REFERENCE TO RELATED APPLICATIONS:
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 617 720 3500
 TELEFAX: 617 720 2441
 INFORMATION FOR SEQUENCING:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 755 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: Double
 POSITION: Linear
 MOLECULE TYPE: cDNA
 RECOMBINATION: No
 ANTI-SENSE: No
 ORIGINAL SOURCE:
 ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: 755
 LOCATION: 51...594
 US 09 791 495 A

Query Match 43.74 Score 437.67 Bk 1 Length 755;
 Best Global Similarity 71.7% Prod. No. 2.50-89;
 Matches 700 Conserved 90 Mismatches 49 Indels 229 Gaps 12

43 581519 755 74
 44 581519 755 60
 45 581519 755 134
 46 581519 755 120
 47 581519 755 154
 48 581519 755 180
 49 581519 755 254
 50 581519 755 240
 51 581519 755 314
 52 581519 755 300

435 581519 755 74
 44 581519 755 60
 45 581519 755 134
 46 581519 755 120
 47 581519 755 154
 48 581519 755 180
 49 581519 755 254
 50 581519 755 240
 51 581519 755 314
 52 581519 755 300

RESULT 6
 US-09-791,495-1
 Sequence 1: Applicant Information
 Patent No. 581519
 GENERAL INFORMATION:
 APPLICANT: Lethy, Bernard
 APPLICANT: Lucus, Sophie
 APPLICANT: de Smet, Charles
 APPLICANT: Godelaine, Danielle
 APPLICANT: Bonn-Falour, Thierry
 TITLE OF INVENTION: Lethy, THUMOR SPECIFIC GENES
 NUMBER OF SEQUENCES: 14
 CORRESPONDENT ADDRESS:
 ADDRESSEE: Well, Greenfield & Sacks, P.C.
 STREET: 600 Atlantic Avenue
 CITY: Boston
 STATE: MA
 COUNTRY: USA
 ZIP: 02210
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 OPERATING SYSTEM: IBM PC compatible
 SOFTWARE: Patent In Release #1.0, Version #1.25

Sequence version 4.5
Copyright (c) 1993 - 2000 CompuLink Ltd.

0M template nucleic search using SW model

Run on: May 1, 2001, 04:00:19 : Search time 329.42 seconds
(without alignments)
24,818 Million cell updates/sec

Hit for: US 09 041 829A-4_copy_1_14
Product score: 14
Sequence: 1 tetra-nucleotide 14

Scoring table: IDENTITY NUC
Gap: 10, 0, 60, 1, 0

Scorched: 6/8276 seqs, 291890651 residues

Total number of hits satisfying chosen parameters: 1556552

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post processing: Minimum Match 0%

Listing first 45 summaries

Database: 1 Nucleic acid 1*

Rank	Score	Match	Length	ID	Description
1	14	100%	1002	19	V50346
2	13	92%	133	19	X12007
3	13	92%	501	21	C51862
4	13	92%	624	20	V89990
5	13	92%	885	10	N90091
6	13	92%	1105	21	C52511
7	13	92%	1864	21	C74482
8	13	92%	6747	14	Q24508
9	12.4	88%	20	18	I78980
10	12.4	88%	43	21	A93915
11	12.4	88%	71	17	I97834

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	14	100%	1002	19	V50346
2	13	92%	133	19	X12007
3	13	92%	501	21	C51862
4	13	92%	624	20	V89990
5	13	92%	885	10	N90091
6	13	92%	1105	21	C52511
7	13	92%	1864	21	C74482
8	13	92%	6747	14	Q24508
9	12.4	88%	20	18	I78980
10	12.4	88%	43	21	A93915
11	12.4	88%	71	17	I97834

C 12	12.4	88.6	71	21	752244	Activated nucleoph
C 13	12.4	88.6	61	21	A76118	Ady filter protein
C 14	12.4	88.6	81	21	A76119	Ady filter protein
C 15	12.4	88.6	84	21	C41994	Human secreted pro
C 16	12.4	88.6	92	12	Q11697	Plasmaid pWBC2.1
C 17	12.4	88.6	106	21	A98817	RNA encoding an as
C 18	12.4	88.6	126	21	A98828	Template DNA enco
C 19	12.4	88.6	166	21	A98829	DNA encoding AV5.1
C 20	12.4	88.6	186	21	C24982	Human secreted pro
C 21	12.4	88.6	190	21	C24982	Human secreted pro
C 22	12.4	88.6	199	21	C10011	Human secreted pro
C 23	12.4	88.6	309	21	A93332	Human colon cancer
C 24	12.4	88.6	310	21	C51941	Arbidopsis thalian
C 25	12.4	88.6	324	16	I26439	Human streptococu
C 26	12.4	88.6	350	17	I07135	Protein-LAL coding
C 27	12.4	88.6	463	21	C56738	Paracryptus grandis
C 28	12.4	88.6	472	14	Q56688	Human brain Expro
C 29	12.4	88.6	481	14	Q66707	Human brain Expro
C 30	12.4	88.6	481	21	C07458	Human secreted pro
C 31	12.4	88.6	487	17	I58552	Streptomyces prot
C 32	12.4	88.6	411	21	C41615	Human secreted pro
C 33	12.4	88.6	415	20	X41299	Human secreted pro
C 34	12.4	88.6	428	21	C01450	Human secreted pro
C 35	12.4	88.6	434	21	C00961	Human secreted pro
C 36	12.4	88.6	444	9	N81993	Human basic fibro
C 37	12.4	88.6	456	21	C42752	Arbidopsis thalian
C 38	12.4	88.6	490	21	C44512	Cell says DNA from
C 39	12.4	88.6	519	17	I07042	Immature DNA from
C 40	12.4	88.6	573	21	F08110	Fusarium venenatu
C 41	12.4	88.6	581	21	A44771	Rat secreted expro
C 42	12.4	88.6	615	21	C55062	Arbidopsis thalian
C 43	12.4	88.6	668	21	C52200	Arbidopsis thalian
C 44	12.4	88.6	712	21	C55062	Arbidopsis thalian
C 45	12.4	88.6	712	21	A96614	Human colon cancer

Acc: 3383385

RESULT 1
V50346
A. Vavab standard, v50346, 1002 bp.
XX V50346:
XX V50346: (Cont. 1/1)
XX V50346: (Cont. 2/1)
XX V50346: (Cont. 3/1)
XX V50346: (Cont. 4/1)
XX V50346: (Cont. 5/1)
XX V50346: (Cont. 6/1)
XX V50346: (Cont. 7/1)
XX V50346: (Cont. 8/1)
XX V50346: (Cont. 9/1)
XX V50346: (Cont. 10/1)
XX V50346: (Cont. 11/1)
XX V50346: (Cont. 12/1)
XX V50346: (Cont. 13/1)
XX V50346: (Cont. 14/1)
XX V50346: (Cont. 15/1)
XX V50346: (Cont. 16/1)
XX V50346: (Cont. 17/1)
XX V50346: (Cont. 18/1)
XX V50346: (Cont. 19/1)
XX V50346: (Cont. 20/1)
XX V50346: (Cont. 21/1)
XX V50346: (Cont. 22/1)
XX V50346: (Cont. 23/1)
XX V50346: (Cont. 24/1)
XX V50346: (Cont. 25/1)
XX V50346: (Cont. 26/1)
XX V50346: (Cont. 27/1)
XX V50346: (Cont. 28/1)
XX V50346: (Cont. 29/1)
XX V50346: (Cont. 30/1)
XX V50346: (Cont. 31/1)
XX V50346: (Cont. 32/1)
XX V50346: (Cont. 33/1)
XX V50346: (Cont. 34/1)
XX V50346: (Cont. 35/1)
XX V50346: (Cont. 36/1)
XX V50346: (Cont. 37/1)
XX V50346: (Cont. 38/1)
XX V50346: (Cont. 39/1)
XX V50346: (Cont. 40/1)
XX V50346: (Cont. 41/1)
XX V50346: (Cont. 42/1)
XX V50346: (Cont. 43/1)
XX V50346: (Cont. 44/1)
XX V50346: (Cont. 45/1)
XX V50346: (Cont. 46/1)
XX V50346: (Cont. 47/1)
XX V50346: (Cont. 48/1)
XX V50346: (Cont. 49/1)
XX V50346: (Cont. 50/1)
XX V50346: (Cont. 51/1)
XX V50346: (Cont. 52/1)
XX V50346: (Cont. 53/1)
XX V50346: (Cont. 54/1)
XX V50346: (Cont. 55/1)
XX V50346: (Cont. 56/1)
XX V50346: (Cont. 57/1)
XX V50346: (Cont. 58/1)
XX V50346: (Cont. 59/1)
XX V50346: (Cont. 60/1)
XX V50346: (Cont. 61/1)
XX V50346: (Cont. 62/1)
XX V50346: (Cont. 63/1)
XX V50346: (Cont. 64/1)
XX V50346: (Cont. 65/1)
XX V50346: (Cont. 66/1)
XX V50346: (Cont. 67/1)
XX V50346: (Cont. 68/1)
XX V50346: (Cont. 69/1)
XX V50346: (Cont. 70/1)
XX V50346: (Cont. 71/1)
XX V50346: (Cont. 72/1)
XX V50346: (Cont. 73/1)
XX V50346: (Cont. 74/1)
XX V50346: (Cont. 75/1)
XX V50346: (Cont. 76/1)
XX V50346: (Cont. 77/1)
XX V50346: (Cont. 78/1)
XX V50346: (Cont. 79/1)
XX V50346: (Cont. 80/1)
XX V50346: (Cont. 81/1)
XX V50346: (Cont. 82/1)
XX V50346: (Cont. 83/1)
XX V50346: (Cont. 84/1)
XX V50346: (Cont. 85/1)
XX V50346: (Cont. 86/1)
XX V50346: (Cont. 87/1)
XX V50346: (Cont. 88/1)
XX V50346: (Cont. 89/1)
XX V50346: (Cont. 90/1)
XX V50346: (Cont. 91/1)
XX V50346: (Cont. 92/1)
XX V50346: (Cont. 93/1)
XX V50346: (Cont. 94/1)
XX V50346: (Cont. 95/1)
XX V50346: (Cont. 96/1)
XX V50346: (Cont. 97/1)
XX V50346: (Cont. 98/1)
XX V50346: (Cont. 99/1)
XX V50346: (Cont. 100/1)
XX V50346: (Cont. 101/1)
XX V50346: (Cont. 102/1)
XX V50346: (Cont. 103/1)
XX V50346: (Cont. 104/1)
XX V50346: (Cont. 105/1)
XX V50346: (Cont. 106/1)
XX V50346: (Cont. 107/1)
XX V50346: (Cont. 108/1)
XX V50346: (Cont. 109/1)
XX V50346: (Cont. 110/1)
XX V50346: (Cont. 111/1)
XX V50346: (Cont. 112/1)
XX V50346: (Cont. 113/1)
XX V50346: (Cont. 114/1)
XX V50346: (Cont. 115/1)
XX V50346: (Cont. 116/1)
XX V50346: (Cont. 117/1)
XX V50346: (Cont. 118/1)
XX V50346: (Cont. 119/1)
XX V50346: (Cont. 120/1)
XX V50346: (Cont. 121/1)
XX V50346: (Cont. 122/1)
XX V50346: (Cont. 123/1)
XX V50346: (Cont. 124/1)
XX V50346: (Cont. 125/1)
XX V50346: (Cont. 126/1)
XX V50346: (Cont. 127/1)
XX V50346: (Cont. 128/1)
XX V50346: (Cont. 129/1)
XX V50346: (Cont. 130/1)
XX V50346: (Cont. 131/1)
XX V50346: (Cont. 132/1)
XX V50346: (Cont. 133/1)
XX V50346: (Cont. 134/1)
XX V50346: (Cont. 135/1)
XX V50346: (Cont. 136/1)
XX V50346: (Cont. 137/1)
XX V50346: (Cont. 138/1)
XX V50346: (Cont. 139/1)
XX V50346: (Cont. 140/1)
XX V50346: (Cont. 141/1)
XX V50346: (Cont. 142/1)
XX V50346: (Cont. 143/1)
XX V50346: (Cont. 144/1)
XX V50346: (Cont. 145/1)
XX V50346: (Cont. 146/1)
XX V50346: (Cont. 147/1)
XX V50346: (Cont. 148/1)
XX V50346: (Cont. 149/1)
XX V50346: (Cont. 150/1)
XX V50346: (Cont. 151/1)
XX V50346: (Cont. 152/1)
XX V50346: (Cont. 153/1)
XX V50346: (Cont. 154/1)
XX V50346: (Cont. 155/1)
XX V50346: (Cont. 156/1)
XX V50346: (Cont. 157/1)
XX V50346: (Cont. 158/1)
XX V50346: (Cont. 159/1)
XX V50346: (Cont. 160/1)
XX V50346: (Cont. 161/1)
XX V50346: (Cont. 162/1)
XX V50346: (Cont. 163/1)
XX V50346: (Cont. 164/1)
XX V50346: (Cont. 165/1)
XX V50346: (Cont. 166/1)
XX V50346: (Cont. 167/1)
XX V50346: (Cont. 168/1)
XX V50346: (Cont. 169/1)
XX V50346: (Cont. 170/1)
XX V50346: (Cont. 171/1)
XX V50346: (Cont. 172/1)
XX V50346: (Cont. 173/1)
XX V50346: (Cont. 174/1)
XX V50346: (Cont. 175/1)
XX V50346: (Cont. 176/1)
XX V50346: (Cont. 177/1)
XX V50346: (Cont. 178/1)
XX V50346: (Cont. 179/1)
XX V50346: (Cont. 180/1)
XX V50346: (Cont. 181/1)
XX V50346: (Cont. 182/1)
XX V50346: (Cont. 183/1)
XX V50346: (Cont. 184/1)
XX V50346: (Cont. 185/1)
XX V50346: (Cont. 186/1)
XX V50346: (Cont. 187/1)
XX V50346: (Cont. 188/1)
XX V50346: (Cont. 189/1)
XX V50346: (Cont. 190/1)
XX V50346: (Cont. 191/1)
XX V50346: (Cont. 192/1)
XX V50346: (Cont. 193/1)
XX V50346: (Cont. 194/1)
XX V50346: (Cont. 195/1)
XX V50346: (Cont. 196/1)
XX V50346: (Cont. 197/1)
XX V50346: (Cont. 198/1)
XX V50346: (Cont. 199/1)
XX V50346: (Cont. 200/1)
XX V50346: (Cont. 201/1)
XX V50346: (Cont. 202/1)
XX V50346: (Cont. 203/1)
XX V50346: (Cont. 204/1)
XX V50346: (Cont. 205/1)
XX V50346: (Cont. 206/1)
XX V50346: (Cont. 207/1)
XX V50346: (Cont. 208/1)
XX V50346: (Cont. 209/1)
XX V50346: (Cont. 210/1)
XX V50346: (Cont. 211/1)
XX V50346: (Cont. 212/1)
XX V50346: (Cont. 213/1)
XX V50346: (Cont. 214/1)
XX V50346: (Cont. 215/1)
XX V50346: (Cont. 216/1)
XX V50346: (Cont. 217/1)
XX V50346: (Cont. 218/1)
XX V50346: (Cont. 219/1)
XX V50346: (Cont. 220/1)
XX V50346: (Cont. 221/1)
XX V50346: (Cont. 222/1)
XX V50346: (Cont. 223/1)
XX V50346: (Cont. 224/1)
XX V50346: (Cont. 225/1)
XX V50346: (Cont. 226/1)
XX V50346: (Cont. 227/1)
XX V50346: (Cont. 228/1)
XX V50346: (Cont. 229/1)
XX V50346: (Cont. 230/1)
XX V50346: (Cont. 231/1)
XX V50346: (Cont. 232/1)
XX V50346: (Cont. 233/1)
XX V50346: (Cont. 234/1)
XX V50346: (Cont. 235/1)
XX V50346: (Cont. 236/1)
XX V50346: (Cont. 237/1)
XX V50346: (Cont. 238/1)
XX V50346: (Cont. 239/1)
XX V50346: (Cont. 240/1)
XX V50346: (Cont. 241/1)
XX V50346: (Cont. 242/1)
XX V50346: (Cont. 243/1)
XX V50346: (Cont. 244/1)
XX V50346: (Cont. 245/1)
XX V50346: (Cont. 246/1)
XX V50346: (Cont. 247/1)
XX V50346: (Cont. 248/1)
XX V50346: (Cont. 249/1)
XX V50346: (Cont. 250/1)
XX V50346: (Cont. 251/1)
XX V50346: (Cont. 252/1)
XX V50346: (Cont. 253/1)
XX V50346: (Cont. 254/1)
XX V50346: (Cont. 255/1)
XX V50346: (Cont. 256/1)
XX V50346: (Cont. 257/1)
XX V50346: (Cont. 258/1)
XX V50346: (Cont. 259/1)
XX V50346: (Cont. 260/1)
XX V50346: (Cont. 261/1)
XX V50346: (Cont. 262/1)
XX V50346: (Cont. 263/1)
XX V50346: (Cont. 264/1)
XX V50346: (Cont. 265/1)
XX V50346: (Cont. 266/1)
XX V50346: (Cont. 267/1)
XX V50346: (Cont. 268/1)
XX V50346: (Cont. 269/1)
XX V50346: (Cont. 270/1)
XX V50346: (Cont. 271/1)
XX V50346: (Cont. 272/1)
XX V50346: (Cont. 273/1)
XX V50346: (Cont. 274/1)
XX V50346: (Cont. 275/1)
XX V50346: (Cont. 276/1)
XX V50346: (Cont. 277/1)
XX V50346: (Cont. 278/1)
XX V50346: (Cont. 279/1)
XX V50346: (Cont. 280/1)
XX V50346: (Cont. 281/1)
XX V50346: (Cont. 282/1)
XX V50346: (Cont. 283/1)
XX V50346: (Cont. 284/1)
XX V50346: (Cont. 285/1)
XX V50346: (Cont. 286/1)
XX V50346: (Cont. 287/1)
XX V50346: (Cont. 288/1)
XX V50346: (Cont. 289/1)
XX V50346: (Cont. 290/1)
XX V50346: (Cont. 291/1)
XX V50346: (Cont. 292/1)
XX V50346: (Cont. 293/1)
XX V50346: (Cont. 294/1)
XX V50346: (Cont. 295/1)
XX V50346: (Cont. 296/1)
XX V50346: (Cont. 297/1)
XX V50346: (Cont. 298/1)
XX V50346: (Cont. 299/1)
XX V50346: (Cont. 300/1)
XX V50346: (Cont. 301/1)
XX V50346: (Cont. 302/1)
XX V50346: (Cont. 303/1)
XX V50346: (Cont. 304/1)
XX V50346: (Cont. 305/1)
XX V50346: (Cont. 306/1)
XX V50346: (Cont. 307/1)
XX V50346: (Cont. 308/1)
XX V50346: (Cont. 309/1)
XX V50346: (Cont. 310/1)
XX V50346: (Cont. 311/1)
XX V50346: (Cont. 312/1)
XX V50346: (Cont. 313/1)
XX V50346: (Cont. 314/1)
XX V50346: (Cont. 315/1)
XX V50346: (Cont. 316/1)
XX V50346: (Cont. 317/1)
XX V50346: (Cont. 318/1)
XX V50346: (Cont. 319/1)
XX V50346: (Cont. 320/1)
XX V50346: (Cont. 321/1)
XX V50346: (Cont. 322/1)
XX V50346: (Cont. 323/1)
XX V50346: (Cont. 324/1)
XX V50346: (Cont. 325/1)
XX V50346: (Cont. 326/1)
XX V50346: (Cont. 327/1)
XX V50346: (Cont. 328/1)
XX V50346: (Cont. 329/1)
XX V50346: (Cont. 330/1)
XX V50346: (Cont. 331/1)
XX V50346: (Cont. 332/1)
XX V50346: (Cont. 333/1)
XX V50346: (Cont. 334/1)
XX V50346: (Cont. 335/1)
XX V50346: (Cont. 336/1)
XX V50346: (Cont. 337/1)
XX V50346: (Cont. 338/1)
XX V50346: (Cont. 339/1)
XX V50346: (Cont. 340/1)
XX V50346: (Cont. 341/1)
XX V50346: (Cont. 342/1)
XX V50346: (Cont. 343/1)
XX V50346: (Cont. 344/1)
XX V50346: (Cont. 345/1)
XX V50346: (Cont. 346/1)
XX V50346: (Cont. 347/1)
XX V50346: (Cont. 348/1)
XX V50346: (Cont. 349/1)
XX V50346: (Cont. 350/1)
XX V50346: (Cont. 351/1)
XX V50346: (Cont. 352/1)
XX V50346: (Cont. 353/1)
XX V50346: (Cont. 354/1)
XX V50346: (Cont. 355/1)
XX V50346: (Cont. 356/1)
XX V50346: (Cont. 357/1)
XX V50346: (Cont. 358/1)
XX V50346: (Cont. 359/1)
XX V50346: (Cont. 360/1)
XX V50346: (Cont. 361/1)
XX V50346: (Cont. 362/1)
XX V50346: (Cont. 363/1)
XX V50346: (Cont. 364/1)
XX V50346: (Cont. 365/1)
XX V50346: (Cont. 366/1)
XX V50346: (Cont. 367/1)
XX V50346: (Cont. 368/1)
XX V50346: (Cont. 369/1)
XX V50346: (Cont. 370/1)
XX V50346: (Cont. 371/1)
XX V50346: (Cont. 372/1)
XX V50346: (Cont. 373/1)
XX V50346: (Cont. 374/1)
XX V50346: (Cont. 375/1)
XX V50346: (Cont. 376/1)
XX V50346: (Cont. 377/1)
XX V50346: (Cont. 378/1)
XX V50346: (Cont. 379/1)
XX V50346: (Cont. 380/1)
XX V50346: (Cont. 381/1)
XX V50346: (Cont. 382/1)
XX V50346: (Cont. 383/1)
XX V50346: (Cont. 384/1)
XX V50346: (Cont. 385/1)
XX V50346: (Cont. 386/1)
XX V50346: (Cont. 387/1)
XX V50346: (Cont. 388/1)
XX V50346: (Cont. 389/1)
XX V50346: (Cont. 390/1)
XX V50346: (Cont. 391/1)
XX V50346: (Cont. 392/1

PR 21-OCT-1999; 990S-0160768.
 PR 21-OCT-1999; 990S-0160770.
 PR 21-OCT-1999; 990S-0160814.
 PR 21-OCT-1999; 990S-0160815.
 PR 22-OCT-1999; 990S-0160980.
 PR 22-OCT-1999; 990S-0160981.
 PR 22-OCT-1999; 990S-0160989.
 PR 25-OCT-1999; 990S-0161404.
 PR 25-OCT-1999; 990S-0161405.
 PR 25-OCT-1999; 990S-0161406.
 PR 26-OCT-1999; 990S-0161359.
 PR 26-OCT-1999; 990S-0161360.
 PR 26-OCT-1999; 990S-0161361.
 PR 28-OCT-1999; 990S-0161920.
 PR 28-OCT-1999; 990S-0161952.
 PR 28-OCT-1999; 990S-0161953.
 PR 29-OCT-1999; 990S-0162142.

Query Match 92.9%; Score 13; DB 21; Length 501;
 Best Local Similarity 100.0%; Pred. No. 4.8e+02;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 cgcctccgcacg 14
 |||||

DB 254 CIGCTCCGCATC 242

RESULT 4
 V8999; 990S-0160768.

ID V89990 standard; cDNA; 624 bp.

AC V89990;

XX 15-FEB-1999 (first entry)

DE EST clone CW707.

XX Human; secreted protein; expressed sequence tag; EST; haematopoiesis;

KW tissue growth; protein; inhibiting; chemotaxis; haematopoiesis; haemostatic;

KW coagulation; fibrinolysis; thrombolytic; anti-inflammatory; cadherin; anti tumour;

XX gene therapy; ss.

OS Homo sapiens.

XX W8984546-A2.

XX 15-OCT-1998.

XX 10-APR-1998; 98W0-0506955.

XX 10-APR-1997; 970S-0838821.

XX (GEM) GENETICS INST INC.

XX Anagnostou M, Jacobs Z, Lavie E, McCoy JM, Murbach D,

XX Racie LA, Spaulding V, Treacy M;

XX W81; 1999-07007706

XX New polynucleotides encoding human secreted proteins - derived from

XX e.g. human blood, kidney, foetal lung, placenta, testes, brain,

XX ovary, pituitary, retina and colon cDNA libraries.

XX Claim 1: Page 396; 618pp; English.

XX The present sequence represents a human expressed sequence tag (EST).

XX the polynucleotide, which is a secreted EST, and the encoded protein

XX are predicted to have useful biological activities which would make

XX them suitable for treating, preventing or ameliorating medical

XX conditions in humans and animals, although no supporting data is

XX given. Suggested activities include nutritional activity, immune

XX stimulating or suppressing activity, haematopoiesis regulating

XX activity, tissue growth activity, anti-inflammation activity,

CC chemotactic/chemokinetic activity, haemostatic and thrombolytic

CC activity, receptor/ligand activity, anti-inflammatory activity,

CC cadherin/tumour invasion suppressor activity, tumour inhibition

CC activity. The polynucleotide may also be useful for gene therapy.

XX Sequence 624 bp; 175 A; 158 C; 129 G; 162 T; 0 other;

Query Match 92.9%; Score 13; DB 25; Length 624;

Best Local Similarity 100.0%; Pred. No. 4.8e+02;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 cgcctccgcacg 14

|||

ID 604 CIGCTCCGCATC 592

RESULT 5

ID N90091 standard; cDNA; 885 bp.

AC N90091;

XX 01-NOV-1989 (first entry)

DE Gene 519 cDNA.

XX cDNA; gene 519; T-cell peptide; functional, activated

KW cells; human; detecting tumours.

XX Homo sapiens.

XX Key Location/Qualifiers

FT CDS 279..666

FT /ftatq a

XX EP320806-A.

XX 21-JUN-1989.

XX 08-DEC-1988; 88EP-0120556.

XX 15-DEC-1987; 87US-0132926.

XX (STRO) LELAND STANFORD JR UNIV.

XX Krensky AM, Davis M, Schall T, Jonstra J;

XX W81; 1986-179693/25.

XX P-PSDB; P90346.

XX New cDNA for gene 519 encoding new T-cell peptide

XX only in functional, activated cells, useful for

XX detecting tumours and assessing T-cell status.

XX Claim: page 8; 12pp; English.

XX cDNA (see corresp. P90346) has the sequence reported for gene 519.

XX It encodes a peptide associated with normal T-cells, and can detect

XX functional, activated T cells to distinguish between neoplastic cells.

XX Sequence 885 bp; 201 A; 260 C; 240 G; 184 T; 0 other;

Query Match 92.9%; Score 13; DB 10; Length 885;

Best Local Similarity 100.0%; Pred. No. 4.9e+02;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 cgcctccgcacg 14

|||

ID 49 cgcctccgcacg 61

CC a consequent frameshift. Transgenic mice can be used as models
 XX of Huntington's disease.

Sequence 20 BP; 5 A; 4 C; 7 G; 4 T; 0 other;

Query Match: 88.6%; Score 12.4; DB 18; Length 20;

Best Local Similarity 92.9%; Pred. No. 9.6e+02;

Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 tetgectccacate 14

11 |1111111111

15 |1111111111

RESULT 10

A94815

10 A94815 standard: DNA; 13 BP.

XX A94815;

XX A94815;

XX 11 JAN-2001 (first entry)

XX DNA encoding an AS linker peptide.

DE Virus vector: melanocyte stimulating hormone receptor; MSH; cytostatic;

XX Virus vector: melanocyte stimulating hormone receptor; MSH; cytostatic;

XX Virus vector: melanocyte stimulating hormone receptor; MSH; cytostatic;

XX Virus vector: melanocyte stimulating hormone receptor; MSH; cytostatic;

XX Virus vector: melanocyte stimulating hormone receptor; MSH; cytostatic;

XX Virus vector: melanocyte stimulating hormone receptor; MSH; cytostatic;

XX Virus vector: melanocyte stimulating hormone receptor; MSH; cytostatic;

XX Virus vector: melanocyte stimulating hormone receptor; MSH; cytostatic;

XX Virus vector: melanocyte stimulating hormone receptor; MSH; cytostatic;

XX Virus vector: melanocyte stimulating hormone receptor; MSH; cytostatic;

XX Virus vector: melanocyte stimulating hormone receptor; MSH; cytostatic;

XX Virus vector: melanocyte stimulating hormone receptor; MSH; cytostatic;

XX Virus vector: melanocyte stimulating hormone receptor; MSH; cytostatic;

XX Virus vector: melanocyte stimulating hormone receptor; MSH; cytostatic;

XX Virus vector: melanocyte stimulating hormone receptor; MSH; cytostatic;

XX Virus vector: melanocyte stimulating hormone receptor; MSH; cytostatic;

XX Virus vector: melanocyte stimulating hormone receptor; MSH; cytostatic;

XX Virus vector: melanocyte stimulating hormone receptor; MSH; cytostatic;

XX Virus vector: melanocyte stimulating hormone receptor; MSH; cytostatic;

XX Virus vector: melanocyte stimulating hormone receptor; MSH; cytostatic;

XX Virus vector: melanocyte stimulating hormone receptor; MSH; cytostatic;

XX Virus vector: melanocyte stimulating hormone receptor; MSH; cytostatic;

XX Virus vector: melanocyte stimulating hormone receptor; MSH; cytostatic;

XX Virus vector: melanocyte stimulating hormone receptor; MSH; cytostatic;

XX Virus vector: melanocyte stimulating hormone receptor; MSH; cytostatic;

XX Virus vector: melanocyte stimulating hormone receptor; MSH; cytostatic;

XX Virus vector: melanocyte stimulating hormone receptor; MSH; cytostatic;

XX Virus vector: melanocyte stimulating hormone receptor; MSH; cytostatic;

XX Virus vector: melanocyte stimulating hormone receptor; MSH; cytostatic;

XX Virus vector: melanocyte stimulating hormone receptor; MSH; cytostatic;

XX Virus vector: melanocyte stimulating hormone receptor; MSH; cytostatic;

XX Virus vector: melanocyte stimulating hormone receptor; MSH; cytostatic;

XX Virus vector: melanocyte stimulating hormone receptor; MSH; cytostatic;

XX Virus vector: melanocyte stimulating hormone receptor; MSH; cytostatic;

XX Virus vector: melanocyte stimulating hormone receptor; MSH; cytostatic;

XX Virus vector: melanocyte stimulating hormone receptor; MSH; cytostatic;

XX Virus vector: melanocyte stimulating hormone receptor; MSH; cytostatic;

XX Virus vector: melanocyte stimulating hormone receptor; MSH; cytostatic;

XX Virus vector: melanocyte stimulating hormone receptor; MSH; cytostatic;

XX Virus vector: melanocyte stimulating hormone receptor; MSH; cytostatic;

XX Virus vector: melanocyte stimulating hormone receptor; MSH; cytostatic;

XX Virus vector: melanocyte stimulating hormone receptor; MSH; cytostatic;

XX Virus vector: melanocyte stimulating hormone receptor; MSH; cytostatic;

XX Virus vector: melanocyte stimulating hormone receptor; MSH; cytostatic;

XX Virus vector: melanocyte stimulating hormone receptor; MSH; cytostatic;

XX Virus vector: melanocyte stimulating hormone receptor; MSH; cytostatic;

XX Virus vector: melanocyte stimulating hormone receptor; MSH; cytostatic;

XX Virus vector: melanocyte stimulating hormone receptor; MSH; cytostatic;

XX Virus vector: melanocyte stimulating hormone receptor; MSH; cytostatic;


```

PN W020004644-AL.
XX
XX 10-AUG 2000.
XX
XX 04-FEB-2000; 2000W0 0802867.
XX
XX 05-FEB 1999; 990S-0118880.
XX
XX (GABR ) GABR RES F00RD.
XX
XX Daniel DL, Prasadgith VR, Emiliotto L, Douglas CL.
XX WPI; 2000-506070/45.
XX
XX Propagation of adenovirus in a host cell, comprising expressing an
XX artificial recombinant receptor on the surface of a host cell and
XX infecting the host with the adenovirus, useful for cell specific gene
XX therapy -
XX
XX Example 8; Page 21; 54pp; English.
XX
XX The present invention relates to a method for propagating adenoviruses
XX in host cells. The method comprises expressing an artificial recombinant
XX receptor on the surface of a host cell and infecting the host with the
XX adenovirus, the virus binding domain of the novel receptor consists of
XX adenovirus serotype 5 (Ad5) fiber protein with a C-terminal His tag
XX (fiber-6His). The present sequence is a coding sequence for fiber-6His
XX peptide. The present sequence was used in the production of vectors
XX encoding the novel receptor of the present invention. The novel vectors
XX are useful for cell specific gene therapy. See B15187 for details of the
XX poly-His tag used to construct the fiber-6His peptide.
XX
XX Sequence 81 BP; 16 A; 12 C; 42 G; 21 T; 0 other;
XX
XX
XX Query Map: 80 69; Score 12.4; DB 21; Length 91;
XX Best Local Similarity 92.9%; Prod. No. 9.7e+02;
XX Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX 07 1 tetacetyccagc 14
XX |||||||||||
XX 78 TTAGCCGCGGATC 65
XX
XX RESULT 15
XX C11994/
XX
XX A: C11994;
XX
XX 06-OCT-2000 (first entry)
XX
XX Human secreted protein 5' EST, SEQ ID NO: 46069.
XX
XX human; 5' EST; expressed sequence tag, secreted protein, cDNA isolation;
XX gene therapy; chromosome mapping; ss.
XX
XX Homo sapiens.
XX
XX EP1034401-A2.
XX
XX 06-SEP-2000.
XX
XX 21-FEB-2000; 2000EP-0200610.
XX
XX 26-FEB-1999; 990S-0122487.
XX
XX (GSE1 ) GENSE1.
XX
XX Thomas Milne Edwards J, Duclert A, Giordano J;
XX WPI; 2000-500381/45.
XX

```

```

XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for
XX obtaining cDNAs and genomic RNAs that correspond to 5' ESTs and for
XX diagnostic, forensic, gene therapy and chromosome mapping procedures -
XX
XX Claim 1; SEQ ID 36069; 71pp; CD-ROM; English.
XX
XX The present sequence is one of a large number of 5' ESTs derived from
XX mRNAs encoding secreted proteins. No ORF has yet been conclusively
XX identified within the present sequence. The 5' ESTs were prepared from
XX total human RNAs or poly(A) RNAs derived from 30 different tissues. EST
XX sequences usually correspond mainly to the 3' untranslated region (UTR)
XX of the mRNA because they are often obtained from oligo dT primed cDNA
XX libraries. Such ESTs are not well suited for isolating cDNA sequences
XX derived from the 5' ends of mRNAs and even in those cases where longer
XX cDNA sequences have been obtained, the full 5' UTR is rarely included.
XX 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be
XX used to obtain full length cDNAs and genomic RNAs. 5' ESTs are also used
XX in diagnostic, forensic, gene therapy and chromosome mapping procedures.
XX They are used to obtain upstream regulatory sequences and to design
XX expression and secretion vectors.
XX
XX Sequence 84 BP; 25 A; 18 C; 46 G; 2 T; 3 other;
XX
XX
XX Query Match: 98.6%; Score 12.4; DB 21; Length 84;
XX Best Local Similarity 92.9%; Prod. No. 9.7e+02;
XX Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX 07 1 tetacetyccagc 14
XX |||||||||||
XX 44 TCTGCGCTCCGCTC 31
XX

```

Search completed: May 1, 2001, 04:00:19
Job time: 7124 sec



Gapcore version 4.5
Copyright (c) 1994 - 2000 CompuGen Ltd.

em nucleic nucleic search, using SW model

Run on: May 1, 2001, 09:03:50, Search time 2512.88 Seconds
(without alignments)
82.165 Million cell updates/sec

Index: us 09_01_829a_4_copy_1_14
Protein source: 14
Sequence: 1 nucleotide/byte 14

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 1204245 swps, 747929632 residues

Total number of hits satisfying chosen parameters: 2566470

Minimum db seq length: 0

Maximum db seq length: 200000000

Post processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: *

1: db_ba1:*

2: db_ba2:*

3: db_ba3:*

4: db_ba4:*

5: db_ba5:*

6: db_ba6:*

7: db_ba7:*

8: db_ba8:*

9: db_ba9:*

10: db_ba10:*

11: db_ba11:*

12: db_ba12:*

13: db_ba13:*

14: db_ba14:*

15: db_ba15:*

16: db_ba16:*

17: db_ba17:*

18: db_ba18:*

19: db_ba19:*

20: db_ba20:*

21: db_ba21:*

22: db_ba22:*

23: db_ba23:*

24: db_ba24:*

25: db_ba25:*

26: db_ba26:*

27: db_ba27:*

28: db_ba28:*

29: db_ba29:*

30: db_ba30:*

31: db_ba31:*

32: db_ba32:*

33: db_ba33:*

34: db_ba34:*

35: db_ba35:*

36: db_ba36:*

37: db_ba37:*

38: db_ba38:*

39: db_ba39:*

40: db_ba40:*

41: db_ba41:*

42: db_ba42:*

43: db_ba43:*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	db	ID	Description
1	14	100.0	224	13	AY017679	oryza sat
2	14	100.0	708	72	AC049025	giardia i
3	14	100.0	795	74	AC066869	giardia i
4	14	100.0	940	72	AC049024	giardia i
5	14	100.0	1002	9	AR042535	AR042535 Sequence
6	14	100.0	1002	92	HSA223040	Homo sapi
7	14	100.0	2649	92	HSA275977	Homo sapi
8	14	100.0	5081	8	AR012720	AR012720 oncorhyn
9	14	100.0	3245	91	H08223094	Homo sapi
10	14	100.0	9859	94	M08F082	L35790 Mus muscula
11	14	100.0	45565	8	FR151119	AL021541 Fugu rnr

```

12 14 100.0 46738 91 HS594110
13 14 100.0 57198 66 AC020115
14 14 100.0 62748 4 AC005417
15 14 100.0 63288 62 AC011762
16 14 100.0 68338 78 AC087440
17 14 100.0 83218 90 AL559880
18 14 100.0 89479 61 AC010444
19 14 100.0 10547 61 AC010009
20 14 100.0 110184 87 AC020906
21 14 100.0 111472 79 AL453145
22 14 100.0 112874 78 AF277415
23 14 100.0 112874 78 AF277415
24 14 100.0 120507 89 AL453558
25 14 100.0 134080 61 AC009523
26 14 100.0 134090 65 AC017815
27 14 100.0 135276 75 AC073603
28 14 100.0 148184 65 AC017777
29 14 100.0 150069 72 AC041042
30 14 100.0 151507 76 AC078802
31 14 100.0 157767 80 AL455344
32 14 100.0 161933 59 AC024704
33 14 100.0 164918 57 AC021599
34 14 100.0 168312 60 AC007484
35 14 100.0 171472 60 AC007400
36 14 100.0 175793 60 AC007401
37 14 100.0 176775 65 AC016978
38 14 100.0 179323 62 AC011852
39 14 100.0 179937 87 AC016683
40 14 100.0 180521 73 AC057817
41 14 100.0 184092 65 AC018715
42 14 100.0 191161 64 AC016043
43 14 100.0 198316 66 AC019157
44 100.0 224744 57 AC022165
45 14 100.0 265512 80 AL455388

```

ALIGNMENTS

```

RESULT 1
LOCUS AY017679 224 bp DNA PLN 07-FEB-2001
DEFINITION Oryza sativa microsatellite M809004 containing (AC)X12, closest to marker R2443, genomic sequence.
ACCESSION AY017679
VERSION 2
KEYWORDS M809004
SOURCE Oryza sativa.
ORGANISM Oryza sativa.
REFERENCE 1 (bases 1 to 224)
AUTHORS Rao N., Barbazuk W.B., Liu J., Wu K. and Barry G.F.
TITLE Simple sequence repeats from Monsanto rice genomic sequences
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 224)
AUTHORS Rao N., Barbazuk W.B., Liu J., Wu K. and Barry G.F.
TITLE Direct Submission
JOURNAL Submitted (10-JAN-2001) Genomics, Monsanto, 800 North Lindbergh Blvd., Creve Coeur, MO 63167, USA
COMMENT derived from rice genomic sequences generated from the Monsanto Rice Genome Sequencing project. Please see http://www.rice-research.org for more information. The sequence data were produced primarily in the laboratories of Dr. Leroy Hood at the University of Washington in Seattle.
FEATURES
source location/Qualifiers
1..224
gb_xref "X060453"
repeat_region 1..224
/locus="microsatellite M809004"
/rpt_unit="tandem

```

```

BASE COUNT 53 a 80 c 52 q 39 t
ORIGIN
AL044582 Human DNA
AC020115 Drosophila
AC005417 Drosophila
AC011762 Drosophila
AC087440 Homo sapi
AL459880 Human DNA
AC010444 Homo sapi
AC010009 Drosophila
AC020906 Homo sapi
AL453145 Homo sapi
AF277415 Homo sapi
AF277415 Homo sapi
AL136358 Human DNA
AC009523 Drosophila
AC017815 Drosophila
AC074603 Mus muscu
AC017777 Drosophila
AC041042 Homo sapi
AC078802 Homo sapi
AL455344 Homo sapi
AC024704 Homo sapi
AC021599 Homo sapi
AC007484 Homo sapi
AC007400 Drosophila
AC007401 Drosophila
AC016978 Homo sapi
AC011852 Homo sapi
AC016683 Homo sapi
AC057817 Homo sapi
AC018715 Homo sapi
AC016043 Homo sapi
AC019157 Homo sapi
AC022165 Homo sapi
AL455388 Homo sapi

```

```

Query Match 100.0% Score 14; DB 14; Length 224;
Best Local Similarity 100.0%; Pred. No. 1.4e+03;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 tctgctccgcatc 14
|||||
Db 78 tctgctccgcatc 91

```

```

RESULT 2
LOCUS AC049025 708 bp DNA HITG 14-APR-2000
DEFINITION Giardia intestinalis clone KJ1363 strain WB-06, LOW-PASS SEQUENCE SAMPLING.
ACCESSION AC049025
VERSION 1
KEYWORDS HITG; PHASE2.
SOURCE Giardia intestinalis.
ORGANISM Giardia intestinalis.
REFERENCE 1 (bases 1 to 708)
AUTHORS Morrison H.G., McArthur A.G., Nixon J., Eskin N.Q., Kim D., Crocker M.C., Hinkle G., Holder M.E. and Sogin M.L.
TITLE Giardia: a model for ancient eukaryotic genome analysis
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 708)
AUTHORS Kim D., Morrison H.G., McArthur A.G., Nixon J., Eskin N.Q., Crocker M.C., Hinkle G., Holder M.E. and Sogin M.L.
TITLE Direct Submission
JOURNAL Submitted (14-APR-2000) Josephine Bay Paul Center for Comparative Molecular Biology and Evolution, Marine Biological Laboratory, 7 MBL Street, Woods Hole, MA 02543-1015, USA
COMMENT * NOTE: This record contains 1 individual
* sequencing reads that have not been assembled into
* contigs. Rups of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.
* 708: contig of 708 bp in length.
FEATURES
source location/Qualifiers
1..708
/organism="Giardia intestinalis"
/strain="WB-06"
/db_xref="taxon:5741"
/clone="KJ1363"
BASE COUNT 217 a 148 c 159 q 187 t 7 others
ORIGIN
1 tctgctccgcatc 14
|||||
1b 560 tctgctccgcatc 567

```

```

RESULT 3
LOCUS AC066869 795 bp DNA HITG 25-APR-2000
DEFINITION Giardia intestinalis clone N10870 strain WB-06, LOW-PASS SEQUENCE SAMPLING.

```


TITLE Direct Submission
JOURNAL Submitted (06-JAN-1998) Lethe B.G., Brussels Branch, Ludwig
 Institute for Cancer Research, 74, avenue Hippocrate, B-1200
 Bruxelles, BELGIUM
REFERENCE 2 (bases 1 to 1002)
AUTHORS Lethe B.G., Lucas S., Michaux F., De Smet C., Godelaine B.,
 Serrano A., De Plaen E. and Boon L.
TITLE LAGE-1, a new gene with tumor specificity
JOURNAL Int. J. Cancer 76 (6), 903-908 (1998)
MEDLINE 98289662
FEATURES Location/Qualifiers
 source
 1..1002
 /organism="Homo sapiens"
 /isolate="individual 1H374"
 /db_xref="taxon:9606"
 /taxon="taxon:9606"
 /cell_line="L8433-MEL"
 /cell_type="melanoma"
 /map="q428"
 /clone="11-1"
 /note="clone 2"
 65..976
 /gene="LAGE-2"
 /gene="LAGE-1"
 65..976
 /genes="LAGE-1"
 /codon_start=1
 /product="LAGE-1 protein"
 /protein_id="CAAL1043.1"
 /db_xref="GI:2255903"
 /translation="MGARSGGPGGPRGPGHGAASAGQRCPCGAPRPSPLQLRLTMDPSSPMEA
 ELVRRILSRVAAFLPEFGAVLKDTIVSGNLLFNSVRQDFEQA:RMFVVEWELGASAP
 EQKARDIPIPKHKVSEQRPGTGTGPPHGAQDCQRCGVAIVNMFSAHLI"
 971..976
 /polyA_site
 993
 /genes="LAGE-1"
 993
 /polyA_site
 179 a 302 c 436 g 185 t
 BASE COUNT
 ORIGIN
 Query Match 100.0% Score 14 DB 92 Length 1002
 Best Local Similarity 100.0% Pred. No. 1,1e+03
 Matches 14 Conserved 0 Mismatches 0 Indels 0 Gaps 0
 97 1 tetactccatccatc 14
 1000 1 TCTGCTCGGCATC 14
 RESULT 7
 HSA275977 2620 bp DNA PK1 08 MAR 2000
 LOCUS Homo sapiens LAGE-2 gene for NV-Pol-1/ANP-2a protein and
 alternative polypeptide LAGE-2alt, exon 1-2.
 ACCESSION A1275977
 VERSION A1275977.1 GI:7208849
 KEYWORDS CIL-defined antigenic peptide; LAGE-2 gene; LAGE-2a; LAGE-2alt;
 human.
 ORGANISM Homo sapiens
 Eukaryota; Chordata; Craniata; Vertebrata; Eut. Therioma;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 (bases 1 to 2630)
AUTHORS De Smet C., Lurquin C., Lethe B., Martelance V. and Boon L.
TITLE DNA methylation is the primary silencing mechanism for a set of
 germ line- and tumor-specific genes with a CpG-rich promoter
JOURNAL Mol. Cell. Biol. 19 (11), 7427-7335 (1999)
MEDLINE 99454989
REFERENCE 2 (bases 1 to 2630)
AUTHORS Lethe B.G.
TITLE Direct Submission
JOURNAL Submitted (03-MAR-2000) Lethe B.G., Molecular Biology, Ludwig
 Institute for Cancer Research, 74, avenue Hippocrate, 1200 -
 Bruxelles, BELGIUM

COMMENT Related sequence AJ275978: rare LAGE 2 mRNA species retaining intron
 2 (LAGE-2b).
FEATURES Location/Qualifiers
 source
 1..2630
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /sex="M"
 /note="gene isolated from human PAC library H911-5"
 <1..907
 /note "CpG island"
 join(908..1240,1913..2047,2287..2577)
 /gene "LAGE-2"
 /note "LAGE-2a form"
 908..1240
 /gene "LAGE-2"
 /number=1
 908..2577
 /gene "LAGE-2"
 1913(972..1240,1913..2047,2287..2425)
 /gene "LAGE-2"
 /note "CIL defined antigenic peptides are processed from
 this product"
 /codon_start=1
 /product "NV-E33 LAGE 2a protein"
 /protein_id "CAH7694.1"
 /db_xref="GI:7208840"
 /translation "MQASERKTCSTCAKRIKICGRIHAGDSNAGAGGAGAGAGGK
 PGCAAPASGSDNAPGPGHGAASAGQRCPCGAPRPSPLQLRLTMDPSSPMEA
 ELAPRSAGAPAPLPEFGAVLKDTIVSGNLLFNTAAHRRQLQISTSTTQQLSLIM
 WIIQFLPLVFLAQPFGGRR"
 1812..1188
 /gene "LAGE-2"
 /codon_start=1
 /product "alternative polypeptide LAGE-2alt"
 /protein_id "CAH7694.1"
 /db_xref="GI:7208841"
 /translation "MLMAQEAFLMAQKAMLAAGERRVPEAEVPAQVQVHRRKE
 EAPRGVRRMAARIQG"
 1241..1912
 /gene "LAGE-2"
 /note "peptide rich (71.74) with quadrupine account ind for
 53.9% of the total"
 /number=1
 1913..2047
 /gene "LAGE-2"
 /number=2
 2048..2286
 /gene "LAGE-2"
 /number=2
 2287..2577
 /gene "LAGE-2"
 /number=4
 2560..2565
 /gene "LAGE-2"
 2577
 /polyA_site
 /gene="LAGE-2"
 448 a 737 c 1026 g 419 t
 BASE COUNT
 ORIGIN
 Query Match 100.0% Score 14 DB 92 Length 2630
 Best Local Similarity 100.0% Pred. No. 9,3e+02
 Matches 14 Conserved 0 Mismatches 0 Indels 0 Gaps 0
 97 1 tetactccatccatc 14
 1000 1 TCTGCTCGGCATC 921
 Db 908 TCTGCTCGGCATC 921
 RESULT 8
 AB012720/c
 LOCUS AB012720 4281 bp mRNA VK1 12 AUG 1998
 DEFINITION Oncorhynchus masou mg120 mRNA for GTP-binding protein, complete


```

Y IADMEAHEDGCPPEWVKTERQFSNFRILNKDGRMDLDETHWTMPDYDHAQAEA
KRIWVSEKREKRLRELELLEWNNMFVSGQATNYGPDLLRNEDEL"
/feature "RCN1"
/number 5
/complement (12621..12703)
/feature "RCN1"
/number 5
/complement (12702..12901)
/feature "RCN1"
/number 5
/complement (12906..13149)
/feature "RCN1"
/number 4
/complement (13148..13208)
/feature "RCN1"
/number 1
/complement (13207..14440)
/feature "RCN1"
/number 3
/feature "ATCG repeat"
/rpt_type "TANDEM"
/complement (14459..14617)
/feature "RCN1"
/number 3
/complement (14616..14686)
/feature "RCN1"
/number 2
/complement (14685..14878)
/feature "RCN1"
/number 2
/complement (14877..16075)
/feature "RCN1"
/number 1
/number 15100
/feature "GI rich minisatellite"
/rpt_unit (15020..15052)
/rpt_type "TANDEM"
/complement (16074..16300)
/feature "RCN1"
/number 1
/number 19440
/feature "GI repeat"
/rpt_type "TANDEM"
/number 20126..20176
/feature "GCT repeat"
/rpt_type "TANDEM"
/number 23416..23527
/feature "A(n) repeat"
/rpt_type "TANDEM"
/number 25635..25756
/feature "15110 shows similarity to human PAX6 genomic
sequence (accession numbers HSCFA15 and HSA1280) across a
number of additional regions detailed below (analysed
Query Match 100.0%; Score 14; DB 8; Length 45565;
Best Local Similarity 100.0%; Pred. No. 5.7e-02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 tetraoctadecate 14
|||||
DB 4068% TGTGGCTGGGATC 40676

RESOLUTION 12
HS594110
LOCUS HS594110 46748 bp DNA PRI 07-FEB-2000
DEFINITION Human DNA sequence from clone RP4-594110 on chromosome 1p43-45.3
contains part of the DSCR11.2 (Down syndrome candidate region 1-like
2) gene, an EST, STSS and GSSs, complete sequence.
ACCESSION AL034582.11 GI:5830348
VERSION AL034582
KEYWORDS hug.
SOURCE human.

```

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Caparrhini; Homnidae; Homo.

1 (bases 1 to 46738)

Baquiley,C.

Direct Submission

Submitted (07-FEB-2000) Sanger Centre, Hinxton, Cambridgeshire,

CR10 1SA, UK. E-mail enquiries: hmcquery@sanger.ac.uk

requests: clonerequest@sanger.ac.uk

On Sep 6, 1999 this sequence was replaced at 15741808.

This sequence has been finished according to sequence map criteria

as follows. An attempt is made to resolve all sequencing problems,

such as compressions and repeats, but not necessarily within known

annotated human repeat sequence elements (e.g. Alu). Where the

sequence is ambiguous, there is an annotation using the 'unsure'

feature key.

During sequence assembly data is compared from overlapping clones.

Where differences are found these are annotated as variations

together with a note of the overlapping clone name. Note that the

variation annotation may not be found in the sequence submission

corresponding to the overlapping clone, as we submit sequences with

only a small overlap as described above.

The following abbreviations are used to associate primary accession

numbers given in the feature table with their source databases:

Em: EMBL; Sw: SWISSPROT; Tr: TrEMBL; Wp: WORMPEP; Information

on the WORMPEP database can be found at

http://www.ebi.ac.uk/Tr/pep/pep2seq/pep2seq.cgi

was generated from part of bacterial clone covering this sequence

chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping

Group. Further information can be found at

http://www.sanger.ac.uk/chr1/chr1

RP4-594110 is from the library RP4-4 constructed at the Roswell

Park Cancer Institute by the group of Peter de Jong. For further

details see http://www.rpi.ac.uk/chr1/chr1

VECTOR: pCYPAC2

IMPORTANT: This sequence is not the entire insert of clone

RP4-594110. It may be shorter because we only sequence overlapping

sections once, or longer because we arrange for a small overlap

between neighbouring submissions.

The true right end of clone RP4-594110 is at 46748 in this

sequence. The true right end of clone RP4-462423 is at 100 in this

sequence.

FEATURES

Location/Qualifiers

1..46738

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="1"

/map="p33-35.4"

/clone_lib="RPCI-4"

/clone="RP4-594110"

/complement(1305..1698)

/note="match: GSS: Em:AQ037179"

/complement(1404..1677)

/note="match: GSS: Em:B16051"

/note="match: GSS: Em:B16051"

/note="match: GSS: Em:B16051"

/note="match: GSS: Em:B16051"

/note="match: GSS: Em:B16051"

/note="match: GSS: Em:B16051"

/note="match: GSS: Em:B16051"

/note="match: GSS: Em:B16051"

/note="match: GSS: Em:B16051"

/note="match: GSS: Em:B16051"

/note="match: GSS: Em:B16051"

/note="match: GSS: Em:B16051"

/note="match: GSS: Em:B16051"

/note="match: GSS: Em:B16051"

/note="match: GSS: Em:B16051"

/note="match: GSS: Em:B16051"

/note="match: GSS: Em:B16051"

/note="match: GSS: Em:B16051"

/note="match: GSS: Em:B16051"

/note="match: GSS: Em:B16051"

/note="match: GSS: Em:B16051"

/note="match: GSS: Em:B16051"

/note="match: GSS: Em:B16051"

/note="match: GSS: Em:B16051"

/note="match: GSS: Em:B16051"

/note="match: GSS: Em:B16051"

/note="match: GSS: Em:B16051"

For further information about this sequence, including its location and relationship to other sequences, please visit our sequence archive web site (<http://www.trinity.org/sequence/>) or send email to bdup@trinity.berkeley.edu.

PI library location: 26-11.

FEATURES

location/qualifiers
1..62748
/organism="Drosophila melanogaster"
/strain="2; cn bw sp"
/db_xref="taxon:7227"
/chromosome="2L"
/map="29F1-29F3"
/clone="115 "PI library, partial SAGE in pMS5821et14AD10"
/size="17,827;11 (3566)"
/note "This clone is a partial bridge extending from tr 1 minimally overlapping its distal neighbor pS06901 (3252) to PI end at bp 62,749."

BASE COUNT: 1925 a 1298 c 1254 g 1796 t

ORIGIN

Query Match 100.0% Score 14; DB 4; Length 62748;
Best Local Similarity 100.0% Ident. No. 5,46,02;
Matches 14; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 1 tetanectocate 14
IIIIIIIIIIII
DB 56376 TCTGCTGGGATC 56489
AC011762 63288 bp DNA HTG 31-JAN-2000
Drosophila melanogaster chromosome 3 clone WAC84901 (D1151)
RCL-98 49.8.1 map 100E-100F strain y; cn bw sp. *** SEQUENCING IN
PROGRESS **, 65 unordered pieces.
AC011762
AC011762.2 GI:6838818
HTG: HUGS_PHASE1
trinity fly
Drosophila melanogaster

Pharyngeal: Neoptera, Endopterygota, Insecta;
Muscomorpha: Ephydroidea: Drosophilidae: Drosophila.
1 (bases 1 to 63288)
Celispor, S. E., Akayama, A., Aradina, T. L., Baxter, E., Blacraj, K. G.,
Butenhof, C., Champe, M., Chavez, C., Chew, M., Ciesiolka, L.,
Boyle, C. M., Fattah, D. E., Gallo, R., George, K. A., Harris, N. L.,
Hinko, A., Hopkins, P. A., Humston, K. A., Humstall, S. P., Karris, K.,
Kearney, L., Lee, B., Lewis, S., Li, P., Lind, H., Mosurel, A. R.,
Nestor, M., Nix, K., Fackel, J. M., Park, S., Pfeiffer, B.,
Richards, S., Sothi, H., Swirskas, P., Wan, F. H., Webster, D.,
Woolley, P., Yang, S., Yeo, M., Yu, C. and Rubin, G. M.
Sequencing of Drosophila melanogaster

Unpublished
2 (bases 1 to 63288)
Celispor, S. E., Akayama, A., Aradina, T. L., Baxter, E., Blacraj, K. G.,
Butenhof, C., Champe, M., Chavez, C., Chew, M., Ciesiolka, L.,
Boyle, C. M., Fattah, D. E., Gallo, R., George, K. A., Harris, N. L.,
Hopkins, P. A., Humston, K. A., Humstall, S. P., Karris, K.,
Kearney, L., Lee, B., Lewis, S., Li, P., Lomolan, M. A., Mazda, P.,
Mosurel, A. R., Nestor, M., Nix, K., Fackel, J. M., Park, S.,
Pfeiffer, B., Pfeiffer, A., Sequiera, A., Sothi, H., Swirskas,
P., Wan, F. H., Webster, D., Yang, S., Yeo, M., Yu, C. and Rubin, G. M.
Robinson, M.
Direct Submission
Submitted (14-07-1999) Drosophila Genome Center, Lawrence Berkeley
Laboratory, MS 64-121, Berkeley, CA 94720, USA
on Jan 31, 2000 this sequence version replaced 91691705.

For further information about this sequence, including its location and relationship to other sequences, please visit our sequence archive web site (<http://www.trinity.org/sequence/>) or send email to bdup@trinity.berkeley.edu. All contigs in this submission were

the following cutoffs: length >= 200 bases.
* NOTE: This is a 'working draft' sequence. It currently consists of 65 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1
658: contig of 608 bp in length
688: gap of unknown length
689
1481: contig of 793 bp in length
1482
1561: gap of unknown length
1562
2344: contig of 703 bp in length
2345
2943: gap of unknown length
2943: contig of 599 bp in length
2944
3023: gap of unknown length
3024
3510: contig of 487 bp in length
3511
3590: gap of unknown length
3591
4001: contig of 411 bp in length
4002
4081: gap of unknown length
4082
4539: contig of 449 bp in length
4531
5268: contig of 658 bp in length
5269
5348: gap of unknown length
5349
5689: contig of 341 bp in length
5690
5770: gap of unknown length
5775: contig of 1006 bp in length
5776
6855: gap of unknown length
6856
7267: contig of 412 bp in length
7268
7347: gap of unknown length
7348
8033: contig of 586 bp in length
8034
8113: gap of unknown length
8114
9116: contig of 1003 bp in length
9117
9196: gap of unknown length
9197
9942: contig of 746 bp in length
9943
10022: gap of unknown length
10023
10608: contig of 586 bp in length
10609
10688: gap of unknown length
10689
11131: contig of 443 bp in length
11132
12211: gap of unknown length
12212
11790: contig of 579 bp in length
11791
11870: gap of unknown length
11871
12565: contig of 695 bp in length
12566
12645: gap of unknown length
12646
13414: contig of 769 bp in length
13415
13494: gap of unknown length
13495
14368: contig of 874 bp in length
14369
14418: gap of unknown length
14419
15088: contig of 640 bp in length
15089
15168: gap of unknown length
15169
16251: contig of 1083 bp in length
16252
16331: gap of unknown length
16332
17722: contig of 1441 bp in length
17723
17852: gap of unknown length
17853
18444: contig of 592 bp in length
18445
18523: gap of unknown length
18525
19607: contig of 1083 bp in length
19608
19687: gap of unknown length
19688
20313: contig of 626 bp in length
20314
20393: gap of unknown length
20394
21212: contig of 819 bp in length
21213
21292: gap of unknown length
21293
22506: contig of 1214 bp in length
22507
22586: gap of unknown length
22587
23231: contig of 635 bp in length
23232
23301: gap of unknown length
23302
24894: contig of 1398 bp in length
24895
24899: gap of unknown length
24899
25632: contig of 743 bp in length
25633
25712: gap of unknown length
25713
27280: contig of 1574 bp in length
27281
27366: gap of unknown length


```
* 27967 28519: contig of 1154 bp in length
* 28520 28593: gap of unknown length
* 28600 30453: contig of 1858 bp in length
* 30458 30538 30547: gap of unknown length
* 30538 32137: contig of 1600 bp in length
* 32137 32217: gap of unknown length
* 32218 34026: contig of 1809 bp in length
* 34027 34106: gap of unknown length
* 34107 36269: contig of 2163 bp in length
* 36270 36349: gap of unknown length
* 36350 38718 38797: contig of 2368 bp in length
* 38718 38797: gap of unknown length
* 38798 40496: contig of 1699 bp in length
* 40497 40576: gap of unknown length
* 40577 42773: contig of 2197 bp in length
* 42774 42853: gap of unknown length
* 42854 44987: contig of 1134 bp in length
* 44988 44987: gap of unknown length
* 44988 47758: contig of 3691 bp in length
* 47759 47838: gap of unknown length
* 47839 48751: contig of 913 bp in length
* 48752 48831: gap of unknown length
* 48832 49498: contig of 667 bp in length
* 49499 49578: gap of unknown length
* 49579 50322: contig of 744 bp in length
* 50323 50402: gap of unknown length
* 50403 50955: contig of 553 bp in length
* 50956 51035: gap of unknown length
* 51036 51530: contig of 495 bp in length
* 51531 51610: gap of unknown length
* 51611 52241: contig of 631 bp in length
* 52242 52321: gap of unknown length
* 52322 52864: contig of 543 bp in length
* 52865 52944: gap of unknown length
* 52945 53398: contig of 454 bp in length
* 53399 53478: gap of unknown length
* 53479 54253: contig of 777 bp in length
* 54254 54333: gap of unknown length
* 54334 54890: contig of 555 bp in length
* 54891 54970: gap of unknown length
* 54971 55570: contig of 600 bp in length
* 55571 55650: gap of unknown length
* 55651 56217: contig of 567 bp in length
* 56218 56297: gap of unknown length
* 56298 56880: contig of 583 bp in length
* 56881 56960: gap of unknown length
* 56961 57530: contig of 576 bp in length
* 57531 57610: gap of unknown length
* 57611 58194: contig of 498 bp in length
* 58195 58760: contig of 566 bp in length
* 58761 58840: gap of unknown length
* 58841 59129: contig of 289 bp in length
* 59130 59209: gap of unknown length
* 59210 59792: contig of 583 bp in length
* 59793 59872: gap of unknown length
* 59873 60440: contig of 568 bp in length
* 60441 60520: gap of unknown length
* 60521 61168: contig of 648 bp in length
* 61169 61248: gap of unknown length
* 61249 61795: contig of 547 bp in length
* 61796 61875: gap of unknown length
* 61876 62431: contig of 556 bp in length
* 62432 62511: gap of unknown length
* 62512 63288: contig of 777 bp in length.
```

FEATURES
source

```
localization/Qualifiers  
1..363288  
/organism "Brosophila melanostator"  
/strain "y; cn bw sp"  
/db_xref "taxon:7223"  
/chromosome "3"  
/map "100E:100F"  
/clone "BACR49B01 (D1151) RPCL 98 49.B.1"
```

```
/clone_lib "RPCL 98 (Roswell Park Cancer Institute  
Brosophila melanostator BAC library, partial Ecoli 10.  
pBACe8.6"  
BASE COUNT 16122 a 14174 c 12761 g 16106 t 5124 others  
ORIGIN  
  
Query Match 100.00; Score 14; DB 62; Length 63288;  
Best Local Similarity 100.00; Pred. No. 5,4002;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
y 1 tetactccacatc 14  
|||||1111111111  
lib 18787 TCTGCTCCGATC 18774  
  
Search completed: May 1, 2001, 04:52:12  
Job time: 11665 sec
```



```

; MOLECULE TYPE: DNA
US-08-894-578-219

Query Match      88.6%; Score 12.4; DB 2; Length 71;
Best Local Similarity 92.9%; Pred. No. 2.9e+02;
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 tetactctccatc 14
    ||||| |||||
DB 22 TTCGCTCCGATC 9

RESULT 7
US-08-248-016-11/c
; Sequence 11, Application US/09248016
; Patent No. 5550109
; GENERAL INFORMATION:
; APPLICANT: Schowetter, Barry S.
; TITLE OF INVENTION: Inducible Defensein Peptide From
; MAMMALIAN EPITHELIA
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; STREET: 1400 I Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3415
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/451,501
; FILING DATE: 26-MAY-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/248,016
; FILING DATE: 24-MAY-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Fordis, Jean H.
; REGISTRATION NUMBER: 30,907
; REFERENCE/JACKET NUMBER: 05387.3017-01000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-408-4000
; TELEFAX: 202-408-4400
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 350 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-451-501-11

Query Match      88.6%; Score 12.4; DB 1; Length 450;
Best Local Similarity 92.9%; Pred. No. 2.9e+02;
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 tetactctccatc 14
    ||||| |||||
DB 258 TTCGCTCCGATC 245

RESULT 9
PCT-US95-06761-11/c
; Sequence 11, Application PC/TUS9506761
; GENERAL INFORMATION:
; APPLICANT: Madadin Pharmaceuticals Inc.
; APPLICANT: 5110 Campus Drive
; APPLICANT: Plymouth Meeting, PA 19462
; TITLE OF INVENTION: Inducible Defensein Peptide From
; MAMMALIAN EPITHELIA
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; STREET: 1400 I Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3415
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent In Release #1.0, Version #1.25

```

```

1  CURRENT APPLICATION DATA:
2  APPLICATION NUMBER: 08/246,985
3  FILING DATE: 20 MAY 1994
4  PRIOR APPLICATION DATA:
5  APPLICATION NUMBER: US 08/248,016
6  FILING DATE: 24 MAY 1994
7  CLASSIFICATION:
8  ATTORNEY/AGENT INFORMATION:
9  NAME: Fabian, Gary R.
10 REGISTRATION NUMBER: 44,876
11 REFERENCE/INVENTOR NUMBER: 08/246,985
12 TELECOMMUNICATION INFORMATION:
13 TELEPHONE: 202-408-4000
14 TELEFAX: 202-408-4400
15 INFORMATION FOR SEQ ID NO: 11:
16 SEQUENCE CHARACTERISTICS:
17 LENGTH: 519 base pairs
18 TYPE: nucleic acid
19 STRANDEDNESS: double
20 TOPOLOGY: linear
21 MOLECULE TYPE: DNA
22 HYDROTHERMAL: NO
23 ORIGINAL SOURCE:
24 INDIVIDUAL ISOLATE: clone D41.2
25 US 08-611-757-48

```

```

1  Query Match: 88.6%; Score 12.4; DB 5; Length 350;
2  Best Local Similarity: 92.0%; Prod. No. 2,900,023;
3  Matches: 13; Conservat 100; 0; Mismatches: 1; Indels: 0; Gaps: 0;
4  QY 1 tetraoctanate 14
5  DB 394 TETRACITRICAC 407

```

```

1  RESULT 11
2  US 08-611-757-48
3  Sequence 48, Application 08/08611757
4  Patent No. 6,552,900
5  GENERAL INFORMATION:
6  APPLICANT: Kim, Judith P.
7  APPLICANT: Boyes, Gregory R.
8  APPLICANT: Waters, John
9  APPLICANT: Clark, E. K.; Clark, Yvonne
10 APPLICANT: Young, Lawrence
11 TITLE OF INVENTION: Reagents and Molecular Cloning Method
12 NUMBER OF SEQUENCES: 106
13 CORRESPONDENCE ADDRESS:
14 ADDRESSEE: Behlinger & Associates
15 STREET: 350 Cambridge Avenue, Suite 250
16 CITY: Palo Alto
17 STATE: CA
18 COUNTRY: USA
19 ZIP: 94306
20 COMPUTER READABLE FORM:
21 MEDIUM TYPE: floppy disk
22 COMPUTER: IBM PC compatible
23 OPERATING SYSTEM: PC DOS/MS DOS
24 SOFTWARE: Behlinger & Associates
25 CURRENT APPLICATION DATA:
26 APPLICATION NUMBER: 08/246,985
27 FILING DATE: 20 MAY 1994
28 CLASSIFICATION:
29 PRIOR APPLICATION DATA:
30 APPLICATION NUMBER: 08/246,985
31 FILING DATE: 20 MAY 1994
32 APPLICATION NUMBER: US 025,496
33 FILING DATE: 24 FEB 1993
34 PRIOR APPLICATION DATA:
35 APPLICATION NUMBER: 08/242,493
36 FILING DATE: 30 JUL 1992
37 ATTORNEY/AGENT INFORMATION:
38 NAME: Fabian, Gary R.
39 REGISTRATION NUMBER: 44,876

```

```

1  TELECOMMUNICATION DATA:
2  APPLICATION NUMBER: 46,000,020
3  TELEPHONE: (415) 324-0880
4  TELEFAX: (415) 324-0960
5  INFORMATION FOR SEQ ID NO: 48:
6  SEQUENCE CHARACTERISTICS:
7  LENGTH: 519 base pairs
8  TYPE: nucleic acid
9  STRANDEDNESS: double
10 TOPOLOGY: linear
11 MOLECULE TYPE: DNA
12 HYDROTHERMAL: NO
13 ORIGINAL SOURCE:
14 INDIVIDUAL ISOLATE: clone D41.2
15 US 08-611-757-48

```

```

1  Query Match: 88.6%; Score 12.4; DB 2; Length 319;
2  Best Local Similarity: 92.0%; Prod. No. 4,000,023;
3  Matches: 13; Conservat 100; 0; Mismatches: 1; Indels: 0; Gaps: 0;

```

```

1  QY 1 tetraoctanate 14
2  DB 394 TETRACITRICAC 407

```

```

1  RESULT 11
2  US 08-611-757-48
3  Sequence 48, Application 08/08611757
4  Patent No. 6,552,900
5  GENERAL INFORMATION:
6  APPLICANT: Kim, Judith P.
7  APPLICANT: Boyes, Gregory R.
8  APPLICANT: Waters, John
9  APPLICANT: Clark, E. K.; Clark, Yvonne
10 APPLICANT: Young, Lawrence
11 TITLE OF INVENTION: Reagents and Molecular Cloning Method
12 NUMBER OF SEQUENCES: 106
13 CORRESPONDENCE ADDRESS:
14 ADDRESSEE: Behlinger & Associates
15 STREET: 350 Cambridge Avenue, Suite 250
16 CITY: Palo Alto
17 STATE: CA
18 COUNTRY: USA
19 ZIP: 94306
20 COMPUTER READABLE FORM:
21 MEDIUM TYPE: floppy disk
22 COMPUTER: IBM PC compatible
23 OPERATING SYSTEM: PC DOS/MS DOS
24 SOFTWARE: Behlinger & Associates
25 CURRENT APPLICATION DATA:
26 APPLICATION NUMBER: 08/246,985
27 FILING DATE: 20 MAY 1994
28 CLASSIFICATION:
29 PRIOR APPLICATION DATA:
30 APPLICATION NUMBER: US 08/246,985
31 FILING DATE: 20 MAY 1994
32 APPLICATION NUMBER: 08/248,016
33 FILING DATE: 24 MAY 1994
34 ATTORNEY/AGENT INFORMATION:
35 NAME: Fabian, Gary R.
36 REGISTRATION NUMBER: 44,876
37 REFERENCE/INVENTOR NUMBER: 46,000,020
38 TELECOMMUNICATION INFORMATION:
39 TELEPHONE: (415) 324-0880
40 TELEFAX: (415) 324-0960
41 INFORMATION FOR SEQ ID NO: 48:
42 SEQUENCE CHARACTERISTICS:
43 LENGTH: 519 base pairs
44 TYPE: nucleic acid
45 STRANDEDNESS: double
46 TOPOLOGY: linear
47 MOLECULE TYPE: DNA
48 HYDROTHERMAL: NO
49 ORIGINAL SOURCE:
50 INDIVIDUAL ISOLATE: clone D41.2
51 US 08-611-757-48

```

Query Match: 88.68; Score 12.4; DB 5; Length 519;
 Best Local Similarity: 92.9%; Pred. No. 30-02;
 Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 tetactctccatc 14
 DB 494 TGTGCTGTCGATC 407

RESULT 12
 US-09-341-829a-4-22/
 Sequence 42, Application US/08946026
 Patent No. 6034218
 GENERAL INFORMATION:
 APPLICANT: Reed, Steven G.
 APPLICANT: Dillon, David C.
 APPLICANT: Iwazaki, Daniel R.
 APPLICANT: Mitcham, Jennifer L.
 TITLE OF INVENTION: COMPOUNDS AND METHODS FOR IMMUNOTHERAPY
 TITLE OF INVENTION: AND IMMUNODIAGNOSIS OF PROSTATE CANCER
 NUMBER OF SEQUENCES: 59
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: SEED AND HERRY LLP
 STREET: 6400 Columbia Center, 701 Fifth Avenue
 CITY: Seattle
 STATE: Washington
 COUNTRY: USA
 ZIP: 98104-7092

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.40

CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/02/946,026
 FILING DATE: 07-OCT-1997
 CLASSIFICATION: 424
 ATTORNEY/AGENT INFORMATION:
 NAME: Makl, David J.
 REGISTRATION NUMBER: 41,492
 REFERENCE/AGENT NUMBER: 210121.424C1
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (206) 622-4900
 TELEFAX: (206) 682-6041
 INFORMATION FOR SEQ ID NO: 32:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 768 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear

US-08-946-026-32
 Query Match: 88.68; Score 12.4; DB 4; Length 768;
 Best Local Similarity: 92.9%; Pred. No. 30-02;
 Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 tetactctccatc 14
 DB 62 TGTGCTGTCGATC 49

RESULT 13
 US-09-475-316A-14
 Sequence 14, Application US/09475216A
 Patent No. 6210942

GENERAL INFORMATION:
 APPLICANT: Lewis, No. 6210942man G.
 APPLICANT: Davin, Laurence B.
 APPLICANT: Linkova-Kostova, Albena T.
 APPLICANT: Fujita, Masayuki
 APPLICANT: Ganq, David R.

APPLICANT: Sarkanen, Simo
 APPLICANT: Ford, Joshua D.
 TITLE OF INVENTION: RECOMBINANT PROTEINS AND METHODS OF USE
 TITLE OF INVENTION: RECOMBINANT PROTEINS AND METHODS OF USE
 FILE REFERENCE: WSUR-1-14793
 CURRENT APPLICATION NUMBER: US/09/475,316A
 CURRENT FILING DATE: 1999-12-40
 PRIOR APPLICATION NUMBER: 09/407,653
 PRIOR FILING DATE: 1999-05-07
 PRIOR APPLICATION NUMBER: PCT/0697/20391
 PRIOR FILING DATE: 1997-11-07
 PRIOR APPLICATION NUMBER: 60/054,480
 PRIOR FILING DATE: 1997-07-31
 PRIOR APPLICATION NUMBER: 60/030,522
 PRIOR FILING DATE: 1996-11-08
 NUMBER OF SEQ ID NOS: 122
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO 14
 LENGTH: 858
 TYPE: DNA
 ORGANISM: Forsythia x intermedia
 FEATURE:
 NAME/KEY: CDS
 LOCATION: (19)..(574)
 US-09-475-316A-14

Query Match: 88.68; Score 12.4; DB 4; Length 858;
 Best Local Similarity: 92.9%; Pred. No. 30-02;
 Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 tetactctccatc 14
 DB 53 tetactctccatc 66

RESULT 14
 US-08-910-731-1
 Sequence 1, Application US/08910731
 Patent No. 5942440
 GENERAL INFORMATION:
 APPLICANT: CHATTERJEE, DEB K.
 APPLICANT: SHANDILYA, HARINI
 TITLE OF INVENTION: Mammalian Ribonuclease Inhibitors and Use Thereof
 NUMBER OF SEQUENCES: 16
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
 STREET: 1100 NEW YORK AVE., N.W., SUITE 600
 CITY: WASHINGTON
 STATE: D.C.
 COUNTRY: USA
 ZIP: 20005-3934

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/910,731
 FILING DATE: (Herewith)
 CLASSIFICATION: 445

PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/795,395
 FILING DATE: 04-FEB-1997
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/794,546
 FILING DATE: 03-FEB-1997
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 60/024,057
 FILING DATE: 16-AUG-1996
 ATTORNEY/AGENT INFORMATION:
 NAME: RSMOND, ROBERT W.
 REGISTRATION NUMBER: 32,894

1 REFERENCE/KEY: NUMBER: 0942-4440003
 2 TELECOMMUNICATION INFORMATION.
 3 TELEPHONE: 202-471-2500
 4 TELEFAX: 202-471-2540
 5 INFORMATION FOR SEQ ID NO: 1:
 6 SEQUENCE CHARACTERISTICS:
 7 LENGTH: 1471 base pairs
 8 TYPE: nucleic acid
 9 STRANDEDNESS: both
 10 TOPOLOGY: both
 11 MOLECULE TYPE: cDNA
 12 FEATURE:
 13 NAME/KEY: CDS
 14 LOCATION: 1..1398
 15 US 09 341 829 A 1

Query Match: 88.6%; Score 12.4; DB 2; Length 1471;
 Best Local Similarity: 92.9%; Prod. No. 40002; Mismatches 1; Indels 0; Gaps 0;
 Matches 14; Conservative 0;

QY 1 tetraocteacate 14
 DB 167 tetraocteacate 180

RESULT 15

US 09 341 829 A 1
 2 Sequence 7, Application US/08910741
 3 Patent No. 5952446

GENERAL INFORMATION

1 APPLICANT: CHATTERJEE, DEB K.

2 APPLICANT: SHANILYA, HARNATI

3 TITLE OF INVENTION: MAMMALIAN Ribonuclease Inhibitors and Use Thereof

4 NUMBER OF SEQUENCES: 15

5 CORRESPONDENCE ADDRESS:

6 ADDRESSEE: STEPHEN KESSLER, GOLDSSTEIN & FOX P.C.L.L.C.

7 STREET: 1100 NEW YORK AVE., N.W., SUITE 600

8 CITY: WASHINGTON

9 STATE: D.C.

10 COUNTRY: USA

11 ZIP: 20005-3934

12 COMPUTER READABLE FORM:

13 MEDIUM TYPE: Floppy disk

14 COMPUTER: IBM PC compatible

15 OPERATING SYSTEM: PC DOS/MS DOS

16 SOFTWARE: Bionin Release #1.0, Version #1.40

17 CURRENT APPLICATION DATA:

18 APPLICATION NUMBER: 09/40919, 741

19 FILING DATE: (Berwick)

20 CLASSIFICATION: 435

21 PRIOR APPLICATION DATA:

22 APPLICATION NUMBER: 09/2795, 495

23 FILING DATE: 04 FEB 1997

24 PRIOR APPLICATION DATA:

25 APPLICATION NUMBER: 09/2794, 546

26 FILING DATE: 04 FEB 1997

27 PRIOR APPLICATION DATA:

28 APPLICATION NUMBER: 60/024, 057

29 FILING DATE: 16 AUG 1996

30 AGENCY/AGENT INFORMATION:

31 NAME: ESMOND, ROBERT W.

32 REGISTRATION NUMBER: 32,893

33 REFERENCE/KEY: NUMBER: 0942-4440003

34 TELECOMMUNICATION INFORMATION:

35 TELEPHONE: 202-471-2500

36 TELEFAX: 202-471-2540

37 INFORMATION FOR SEQ ID NO: 7:

38 SEQUENCE CHARACTERISTICS:

39 LENGTH: 1471 base pairs

40 TYPE: nucleic acid

41 STRANDEDNESS: both

42 TOPOLOGY: both

1 MOLECULE TYPE: cDNA
 2 US 09 341 829 A 1

Query Match: 88.6%; Score 12.4; DB 2; Length 1471;
 Best Local Similarity: 92.9%; Prod. No. 40002; Mismatches 1; Indels 0; Gaps 0;
 Matches 14; Conservative 0;

QY 1 tetraocteacate 14
 DB 167 tetraocteacate 180

Search completed: May 1, 2001, 09:54:26
 Job time: 7276 sec

Copyright (c) 1993 - 2000 CompuLink Ltd.

bioWare version 4.5

CM nucleotide nucleotide search, using sw mode

Run date: May 1, 2001, 09:04:53, Search time 4766.75 Seconds
(without alignments)
42,298 Million cell updates/sec

Title: US 09-341-829A-4 copy_1_14
Posttest score: 14
Sequence: 1 restriction enzyme 14

Scoring table: IDENTIFY NUC
Gapop 19.0, Gapext 1.0

Searches: 50,200 17 steps, 46,904,000 reads

Total number of hits satisfying chosen parameters: 19247044

Minimum hit seq length: 0
Maximum hit seq length: 200000000

Post processing: Minimum Match 0%
Maximum Match 100%
First and Last 45 summaries

Database:

EST:
1: qb_ost11:
2: qb_ost2:
3: qb_ost3:
4: qb_ost4:
5: qb_ost5:
6: qb_ost6:
7: qb_ost7:
8: qb_ost8:
9: qb_ost9:
10: qb_ost10:
11: qb_ost11:
12: qb_ost12:
13: qb_ost13:
14: qb_ost14:
15: qb_ost15:
16: qb_ost16:
17: qb_ost17:
18: qb_ost18:
19: qb_ost19:
20: qb_ost20:
21: qb_ost21:
22: qb_ost22:
23: qb_ost23:
24: qb_ost24:
25: qb_ost25:
26: qb_ost26:
27: qb_ost27:
28: qb_ost28:
29: qb_ost29:
30: qb_ost30:
31: qb_ost31:
32: qb_ost32:
33: qb_ost33:
34: qb_ost34:
35: qb_ost35:
36: qb_ost36:
37: qb_ost37:
38: qb_ost38:
39: qb_ost39:
40: qb_ost40:
41: qb_ost41:
42: qb_ost42:
43: qb_ost43:
44: qb_ost44:
45: qb_ost45:
46: qb_ost46:
47: qb_ost47:

44: em_osthum10:
45: em_osthum11:
46: em_osthum12:
47: em_osthum13:
48: em_osthum14:
49: em_osthum15:
50: em_osthum16:
51: em_osthum17:
52: em_osthum18:
53: em_osthum19:
54: em_osthum20:
55: em_osthum21:
56: em_osthum22:
57: em_osthum23:
58: em_osthum24:
59: em_osthum25:
60: em_osthum26:
61: em_osthum27:
62: em_osthum28:
63: em_osthum29:
64: em_osthum30:
65: em_osthum31:
66: em_osthum32:
67: em_osthum33:
68: em_osthum34:
69: em_osthum35:
70: em_osthum36:
71: em_osthum37:
72: em_osthum38:
73: em_osthum39:
74: em_osthum40:
75: em_osthum41:
76: em_osthum42:
77: em_osthum43:
78: em_osthum44:
79: em_osthum45:
80: em_osthum46:
81: em_osthum47:
82: em_osthum48:
83: em_osthum49:
84: em_osthum50:
85: em_osthum51:
86: em_osthum52:
87: em_osthum53:
88: em_osthum54:
89: em_osthum55:
90: em_osthum56:
91: em_osthum57:
92: em_osthum58:
93: em_osthum59:
94: em_osthum60:
95: em_osthum61:
96: em_osthum62:
97: em_osthum63:
98: em_osthum64:
99: em_osthum65:
100: em_osthum66:
101: em_osthum67:
102: qb_ost25:
103: qb_ost26:
104: qb_ost27:
105: qb_ost28:
106: qb_ost29:
107: qb_ost30:
108: qb_ost31:
109: qb_ost32:
110: qb_ost33:
111: qb_ost34:
112: qb_ost35:
113: qb_ost36:
114: qb_ost37:
115: qb_ost38:
116: qb_ost39:

117: qb_est48:*
 118: qb_est49:*
 119: qb_est50:*
 120: qb_est51:*
 121: qb_est52:*
 122: qb_est53:*
 123: qb_est54:*
 124: qb_est55:*
 125: qb_est56:*
 126: qb_est57:*
 127: qb_est58:*
 128: qb_est59:*
 129: qb_est60:*
 130: qb_est61:*
 131: qb_est62:*
 132: qb_est63:*
 133: qb_est64:*
 134: qb_est65:*
 135: qb_est66:*
 136: qb_est67:*
 137: qb_est68:*
 138: qb_est69:*
 139: qb_est70:*
 140: qb_est71:*
 141: qb_est72:*
 142: qb_est73:*
 143: qb_est74:*
 144: qb_est75:*
 145: qb_est76:*
 146: qb_est77:*
 147: qb_est78:*
 148: qb_est79:*
 149: qb_est80:*
 150: qb_est81:*
 151: qb_est82:*
 152: qb_est83:*
 153: qb_est84:*
 154: qb_est85:*
 155: qb_est86:*
 156: qb_est87:*
 157: qb_est88:*
 158: qb_est89:*
 159: qb_est90:*
 160: qb_est91:*
 161: qb_est92:*
 162: qb_est93:*
 163: qb_est94:*
 164: qb_est95:*
 165: qb_est96:*
 166: qb_est97:*
 167: qb_est98:*
 168: qb_est99:*
 169: qb_est100:*
 170: qb_est101:*
 171: qb_est102:*
 172: qb_est103:*
 173: qb_est104:*
 174: qb_est105:*
 175: qb_est106:*
 176: qb_est107:*
 177: qb_est108:*
 178: qb_est109:*
 179: qb_est110:*
 180: qb_est111:*
 181: qb_est112:*
 182: qb_est113:*
 183: qb_est114:*
 184: qb_est115:*
 185: qb_est116:*
 186: qb_est117:*
 187: qb_est118:*
 188: qb_est119:*
 189: qb_est120:*

190: em_qss_pln1:*
 191: em_qss_pln2:*
 192: em_qss_pro:*
 193: em_qss_rod1:*
 194: em_qss_rod2:*
 195: em_qss_rod3:*
 196: em_qss_rod4:*
 197: em_qss_rod5:*
 198: em_qss_vrt1:*
 199: em_qss_vrt2:*
 200: em_qss_vrt3:*
 201: qb_qss1:*
 202: qb_qss2:*
 203: qb_qss3:*
 204: qb_qss4:*
 205: qb_qss5:*
 206: qb_qss6:*
 207: qb_qss7:*
 208: qb_qss8:*
 209: qb_qss9:*
 210: qb_qss10:*
 211: qb_qss11:*
 212: qb_qss12:*
 213: qb_qss13:*
 214: qb_qss14:*
 215: qb_qss15:*
 216: qb_qss16:*
 217: qb_qss17:*
 218: qb_qss18:*
 219: qb_qss19:*
 220: qb_qss20:*
 221: qb_qss21:*
 222: qb_qss22:*
 223: qb_qss23:*
 224: qb_qss24:*
 225: qb_qss25:*
 226: qb_qss26:*
 227: qb_qss27:*
 228: qb_qss28:*
 229: qb_qss29:*
 230: qb_qss30:*
 231: qb_qss31:*
 232: qb_qss32:*
 233: qb_qss33:*
 234: qb_qss34:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result	No.	Score	Query	Match	Length	DB	ID	Description
1	14	100.0	168	117	AW579092			AW579092 RCO C1046
2	14	100.0	292	124	BB055977			BB055977 BB055977
3	14	100.0	402	106	AB040252			AB040252 AB040252
4	14	100.0	429	104	AL282204			AL282204 AL282204
5	14	100.0	472	3	AA155178			AA155178 mtr1107.1
6	14	100.0	478	165	BE346096			BE346096 SP19406.Y
7	14	100.0	525	166	HE367560			HE367560 P11.9_E03
8	14	100.0	541	215	AZ068067			AZ068067 RPT1_23-3
9	14	100.0	545	174	BC084012			BC084012 B3095E04
10	14	100.0	548	2	AA141767			AA141767 CR02398.C
11	14	100.0	554	219	AZ221492			AZ221492 DM0342103
12	14	100.0	570	172	BC071181			BC071181 B3095E04
13	14	100.0	597	18	AI294860			AI294860 L008430.5
14	14	100.0	616	225	AZ655079			AZ655079 L03-55R17
15	14	100.0	618	232	FR0041522			AL129014 Funct. r001
16	14	100.0	619	242	FR0042367			AL129079 Funct. r001
17	14	100.0	623	22	AI601278			AI601278 L009405.X
18	14	100.0	638	117	AW611303			AW611303 mtr06009.Y

/clone_lib "wajfen full-length enriched, 12 days embryo
 male wolffian duct"
 /sex "male"
 /issue_type "welfian duct includes surrounding region"
 /dev_stage "12 days embryo"
 /lab_host "b10B"
 /note "site_1: SalI; site_2: BamHI; cDNA library was
 prepared and sequenced in Mouse Genome Encyclopedia
 Project of Genome Exploration Research Group in NIKOM
 Genomic Sciences Center and Genome Science Laboratory in
 RIKEN, Division of Experimental Animal Research in Riken
 contributed to prepare mouse tissues 1st strand cDNA was
 primed with a primer [5'
 CAGAGAGAGAGATGACAGAGCTGTTTTTTTTTTTTTTT 3'], cDNA was
 prepared by using probalase thermo-activated reverse
 transcriptase and subsequently enriched for full-length by
 cap-trapper. Second strand cDNA was prepared with the
 primer adaptor of sequence [5'
 CAGAGAGAGATTCGACTTAATTAATTCGCGCGCGCGC 3']. cDNA
 was cloned into the xbaI and BamHI sites. Vector: a
 modified pluescript KS(+) after bulk excision from lambda
 FIC 1. Cloning sites, 5' end: SalI; 3' end: BamHI"
 BASE COUNT 63 a 68 c 56 g 105 t
 ORIGIN
 1 tetgctccgacatc 14
 |||||||
 db 43 tctgctccgacatc 46

Query Match 100.0% Score 14; DB 124; Length 292;
 Best Local Similarity 100.0%; Pred. No. 2.5e+03;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 tctgctccgacatc 14
 |||||||
 DB 43 tctgctccgacatc 46

RESULT 3
 LOCUS A0040252/c
 DEFINITION A0040252 Mouse four-cell-embryo cDNA Mus musculus cDNA clone
 J0802503 3', mRNA sequence.
 ACCESSION A0040252
 VERSION A0040252.1 GI:3954072
 KEYWORDS EST
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 402)
 K.O.M.S.H., Kitchen, J.K., Ward, X., Ibrat, J.A., Sun, T., DePalma, C.E.,
 Liang, Y., Kurul, G., Sharara, R., Lim, M.K. and Doi H
 Systematic analyses of genes expressed in 4-cell mouse embryo (The
 EPAT/DOE Project at Wayne State University)
 Contact: Hirotsumi Doi
 Doi Biosymetry Project, PEATO
 Japan Science and Technology Corporation (JST)
 Wako Marine East 12F, 2-8 Nakase, Mihama-ku, Chiba 261-71, Japan
 Email: hdoibio@jst.go.jp.
 Location/Qualifiers
 1..402
 /organism "Mus musculus"
 /strain "w57H/6J"
 /DB_Xref "taxon:10090"
 /clone "J0802503"
 /note "lib="mouse four-cell embryo cDNA"
 /dev_stage "four-cell-embryo"
 141 a 76 c 64 g 120 t 1 others
 BASE COUNT
 ORIGIN
 Query Match 100.0% Score 14; DB 136; Length 402;
 Best Local Similarity 100.0%; Pred. No. 2.5e+03;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 tctgctccgacatc 14
 |||||||
 DB 233 tctgctccgacatc 220

RESULT 4
 LOCUS A2282204/c
 DEFINITION 4A3A AAG-B 05 R Anopheles gambiae immune competent 4A3A Anophelies
 gambiae cDNA clone 4A3A AAG-B 05, mRNA sequence.
 ACCESSION A2282204
 VERSION A2282204.1 GI:6930083
 KEYWORDS EST
 SOURCE African malaria mosquito.
 ORGANISM Anopheles gambiae
 Eukaryota; Metazoa; Arthropoda; Insecta; Diptera; Nematocera; Culicoidae;
 Pteroptera; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae
 1 (bases 1 to 439)
 Dimopoulos, G., Casavant, T.L., Chant, S., Schreier, L., Roberts, J.,
 Donohue, M., Schultz, J., Jones, V., Bork, P., Ansorge, W., Soares, H. B.
 and Kafatos, F. G.
 Anopheles gambiae pilot gene discovery project: identification of
 mosquito innate immunity genes from expressed sequence tags
 generated from immune-competent cell lines
 Proc. Natl. Acad. Sci. U.S.A. 97 (14), 9619-9624 (2000)
 20300950
 Contact: Dimopoulos G
 Foris G. Kafatos Laboratory
 European Molecular Biology Laboratory
 Meyerhofstrasse 1, 69117 Heidelberg, Germany.
 Location/Qualifiers
 1..439
 /organism "Anopheles gambiae"
 /strain "4A 1/1"
 /DB_Xref "taxon:7165"
 /clone "4A3A-AAG-B-09"
 /clone_lib "Anopheles gambiae immune competent 4A3A"
 /cell_line "immune competent 4A3A"
 /lab_host "E. coli DH10B"
 /note="Vector: pTZ19-Pac (Pharmacia) with a modified
 polylinker; Site_1: EcoRI; Site_2: NotI; sequenced from
 forward priming site which reads from the 3' end of the
 cDNA. The 4A3A is a directionally cloned and normalized
 cDNA library that was constructed from the 4A3A cell line
 oligo-1 primed cDNA according to: Ronaldo, Lourenco & Soares
 (1996) : Normalization and Subtraction: Two approaches to
 facilitate Gene Discovery, Genome Research 6, 791-800."

BASE COUNT 144 a 112 c 122 g 61 t
 ORIGIN
 Query Match 100.0% Score 14; DB 104; Length 439;
 Best Local Similarity 100.0%; Pred. No. 2.5e+04;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 tctgctccgacatc 14
 |||||||
 DB 483 tctgctccgacatc 370

RESULT 5
 LOCUS AA155178/c
 DEFINITION T801197.11 Statobates musae heart (4947376) Mus musculus (ERA clone)
 IMAGE:603877 5' similar to AB003272 FIBRILLIN 2 (HUMAN); AB022494
 Mus musculus fibriclin-1 (MUSE); mRNA sequence.
 ACCESSION AA155178
 VERSION AA155178.1 GI:1726802
 KEYWORDS EST
 SOURCE house mouse.
 ORGANISM Mus musculus

with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent *E. coli* XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 167 a 100 c 111 q 176 t
ORIGIN
Query Match 100.0%; Score 14; DB 219; Length 554;
Best Local Similarity 100.0%; Pred. No. 2.5e+03;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 tetgctccgcacatc 14
|||||
DB 495 TGTGCTCCGCATC 508

RESULT 12
LOCUS B0071181/c 570 bp mRNA EST 26-JAN-2001
DEFINITION H3095F04 3' NIA Mouse 15K cDNA Clone Set Mus musculus cDNA clone
H3095F04 3' mRNA sequence.
ACCESSION B0071181
VERSION B0071181.1 GI:12554750
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 570)
REFERENCE Karqui, G.J., Indekula, D.B., Qian, Y., Lim, M.K., Jaradat, S.A., Tanaka
J.L.S., Carter, M.G., and Ko, M.S.H.
Verification and initial annotation of NIA mouse 15K cDNA clone set
Unpublished (2001)
OTHER ESTs: H3095F04-5
Contact: George J. Karqui
Laboratory of Genetics
National Institute on Aging/National Institutes of Health
345 Cassell Drive, Suite 400, Baltimore, MD 21224-6920, USA
Email: cdna@nslu-grc.nia.nih.gov
This clone set has been freely distributed to the community. Please
visit <http://nslu-grc.nia.nih.gov/15k.htm> for details.
Plate: H3095 row: F column: 04
Seq primer: -21M13 forward
High quality sequence stop: 570
POLYA Yes.
FEATURES
source
1..570
Location/Qualifiers
/organism="Mus musculus"
/strain="c57BL/6J"
/db_xref="taxon:10090"
/clone="H3095F04"
/clone_lib="NIA Mouse 15K cDNA clone set"
/sex="clones arrayed from a mixture of cDNA libraries"
/low_string="clones arrayed from a mixture of cDNA libraries"
/lab_host="DH10B"
/note="Vector: pSPORT1; Site_1: SalI; Site_2: NotI; This clone is among a rearrayed set of 15,247 clones from 11 embryo cDNA libraries (including preimplantation stage embryos from unfertilized egg to blastocyst, embryonic part of E7.5 embryo, extraembryonic part of E7.5 embryo, and E12.5 female mesonephros/ovad) and one newborn cDNA library. Average insert size 1.5 kb. All source libraries are cloned unidirectionally with Oligo(dT)-Not primers. References include: (1) Genome-wide expression profiling of mid-gestation placenta and embryo using a 15,000 mouse developmental cDNA microarray. 2000, Proc. Natl. Acad. Sci. U.S.A. 97: 9127-9132; (2) Large-scale cDNA analysis reveals phased gene expression patterns during preimplantation mouse development, 2000, Development, 127: 1747-1749; (3) Genome-wide mapping of

unselected transcripts from extraembryonic tissue of 7.5-day mouse embryos reveals enrichment in the 1st complex and under-representation on the X chromosome, 1998, Hum Mol Genet 7: 1967-1978."

BASE COUNT 148 a 122 c 92 q 168 t
ORIGIN
Query Match 100.0%; Score 14; DB 172; Length 570;
Best Local Similarity 100.0%; Pred. No. 2.5e+03;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 tetgctccgcacatc 14
|||||
DB 233 TGTGCTCCGCATC 220

RESULT 13
LOCUS A1294860 597 bp mRNA EST 01-DEC-1998
DEFINITION LP08330 5prime 1P Drosophila melanogaster larval-early pupal poly2
Drosophila melanogaster cDNA clone LP08330 5prime. mRNA sequence.
ACCESSION A1294860
VERSION A1294860.1 GI:3944267
KEYWORDS EST.
SOURCE fruit fly.
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Insecta; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 597)
REFERENCE Harvey, D., Hovav, L., Evans-Bach, M., Pendleton, J., Su, C., Blockstein
P., Lewis, S., and Rubin, G.M.
Black/HHMI Drosophila EST Project
Unpublished (1997)
Contact: Harvey, D.
G. M. Rubin-Molecular and Cell Biology
University of California Berkeley
539 LSA, Berkeley, CA 94720-3200, USA
Fax: 510 643 9947
Email: http://www.fruitfly.org/est_est-struct/lyp08330
Plate: 84 row: C column: 6
High quality sequence stop: 443.
FEATURES
source
1..597
Location/Qualifiers
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone_lib="1P Drosophila melanogaster larval-early pupal poly2"
/sex="male and female"
/low_string="larval-pupae"
/lab_host="DH5-alpha"
/note="Organ: whole body; Vector: poly2; Site_1: EcoRI; Site_2: XbaI; Sized fractionated cDNAs were directly ligated into poly2. Plasmid cDNA library."

BASE COUNT 133 a 143 c 181 q 140 t
ORIGIN
Query Match 100.0%; Score 14; DB 18; Length 597;
Best Local Similarity 100.0%; Pred. No. 2.6e+03;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 tetgctccgcacatc 14
|||||
DB 334 TGTGCTCCGCATC 347

RESULT 14
LOCUS AZ465079 616 bp DNA GSS 14 DEC-2000

GenCorp version 4.1b
Copyright (c) 1993-2000 Compugen Ltd.

QM nucleotide nucleotide search, using sw model

Run on: May 1, 2001, 09:52:12 : Search time 25:22.68 Seconds
(without alignments)
2969,667 Million cell updates/sec

131 locs QS 09 341 829A 4 copy_250_755
for loc 350, loc 506
Sequences 1 dependent query seqs, created at cctccaa 506

Search table: IDENTITY NIP
Gapop 10.0 : Gapex 1.0

Searches: 1283235 seqs, 737892952 residues 2566A70
Total number of hits satisfying chosen parameters:

Minimum hit seq length: 6
Maximum hit seq length: 20000000

Post processing: Minimum Match: 0K
Maximum Match: 100K
Fastest First: 45 summaries

Database: GenBank:
1: db_ba1:
2: db_ba2:
3: db_ba3:
4: db_ba4:
5: db_ba5:
6: db_ba6:
7: db_ba7:
8: db_ba8:
9: db_ba9:
10: db_ba10:
11: db_ba11:
12: db_ba12:
13: db_ba13:
14: db_ba14:
15: db_ba15:
16: db_ba16:
17: db_ba17:
18: db_ba18:
19: db_ba19:
20: db_ba20:
21: db_ba21:
22: db_ba22:
23: db_ba23:
24: db_ba24:
25: db_ba25:
26: db_ba26:
27: db_ba27:
28: db_ba28:
29: db_ba29:
30: db_ba30:
31: db_ba31:
32: db_ba32:
33: db_ba33:
34: db_ba34:
35: db_ba35:
36: db_ba36:
37: db_ba37:
38: db_ba38:
39: db_ba39:
40: db_ba40:
41: db_ba41:
42: db_ba42:
43: db_ba43:

44: cm_ov:
45: cm_pat:
46: cm_ph:
47: cm_pl:
48: cm_ro:
49: cm_sfs:
50: cm_sy:
51: cm_un:
52: cm_v1:
53: db_sfs1:
54: db_sfs2:
55: db_sfs3:
56: db_sy:
57: db_un:
58: db_v1:
59: db_v12:
60: db_ha1:
61: db_ha2:
62: db_ha3:
63: db_ha4:
64: db_ha5:
65: db_ha6:
66: db_ha7:
67: db_ha8:
68: db_ha9:
69: db_ha10:
70: db_ha11:
71: db_ha12:
72: db_ha13:
73: db_ha14:
74: db_ha15:
75: db_ha16:
76: db_ha17:
77: db_ha18:
78: db_ha19:
79: db_ha20:
80: db_ha21:
81: db_ha22:
82: db_ha23:
83: db_ha24:
84: db_ha25:
85: db_ha26:
86: db_ha27:
87: db_ha28:
88: db_ha29:
89: db_ha30:
90: db_ha31:
91: db_ha32:
92: db_ha33:
93: db_ha34:
94: db_ha35:
95: db_ha36:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	506	100.0	993	9	AX024089	AX024089 Sequence
2	506	100.0	993	92	HS0012835	AJ012835 Homo sapi
3	506	100.0	1002	9	AB042533	AB042533 Sequence
4	506	100.0	1002	92	HS0220040	AJ220040 Homo sapi
5	422	83.4	112874	78	AF277315	AF277315 Homo sapi
6	418.8	82.3	9245	91	HS223093	AJ223093 Homo sapi
7	465.6	72.3	873	92	HS0275978	AJ275978 Homo sapi
8	405.2	60.3	2630	92	HS0275977	AJ275977 Homo sapi
9	405.2	60.3	112874	78	AF277315	AF277315 Homo sapi
10	219	43.3	679	9	AX024685	AX024685 Sequence
11	219	43.3	679	92	HS0012833	AJ012833 Homo sapi

[illegible]

[illegible]


```

Query Match          100.0%; Score 506; DB 1; Length 1002;
Best Local Similarity 100.0%; Pred. No. 1.9e-127;
Matches 506; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 aaatcccatgagcgtgagcgtctctgacagagatggaatgacccatgagagcccaaaag 60
|||||
DB 250 GGGTCGGATGGGATGAGTCTCTGGGAGATGGAAAGGTGTCCTGGGGGCCAGAGAG 409
|||||

QY 61 acacacacacacacacacacacacacacacacacacacacacacacacacacacacac 120
|||||
DB 310 GAGACACACACACACACACACACACACACACACACACACACACACACACACACACAC 469
|||||

QY 121 acacacacacacacacacacacacacacacacacacacacacacacacacacacacac 180
|||||
DB 370 GAGACACACACACACACACACACACACACACACACACACACACACACACACACACAC 429
|||||

QY 181 tctaaacacacacacacacacacacacacacacacacacacacacacacacacacacacac 240
|||||
DB 430 TCTGAAGACACTTACCGGCTGGGAGGACACACACACACACACACACACACACACACAC 489
|||||

QY 241 aaacacacacacacacacacacacacacacacacacacacacacacacacacacacacac 300
|||||
DB 490 GCAAGGCGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 549
|||||

QY 301 acacacacacacacacacacacacacacacacacacacacacacacacacacacacacac 360
|||||
DB 550 GCAAGGCGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 609
|||||

QY 361 acacacacacacacacacacacacacacacacacacacacacacacacacacacacacac 420
|||||
DB 610 ACACACACACACACACACACACACACACACACACACACACACACACACACACACACACAC 669
|||||

QY 421 tctaaacacacacacacacacacacacacacacacacacacacacacacacacacacacac 480
|||||
DB 670 TGATGATGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 729
|||||

QY 481 acacacacacacacacacacacacacacacacacacacacacacacacacacacacacac 506
|||||
DB 730 AGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 755
|||||

```

```

RESULT 2
US-08-791-495-6
: Sequence 6; Application US/08791495
: Patent No. 5811519
: GENERAL INFORMATION:
: APPLICANT: Leth, Bernard
: APPLICANT: Lucas, Sophie
: APPLICANT: De Smet, Charles
: APPLICANT: Godelaine, Danielle
: APPLICANT: Boon-Fallour, Thierry
: TITLE OF INVENTION: LL-1 TUMOR SPECIFIC GENES
: NUMBER OF SEQUENCES: 14
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Wolf, Greenfield & Sacks, P.C.
: STREET: 600 Atlantic Avenue
: CITY: Boston
: STATE: MA
: COUNTRY: USA
: ZIP: 02210
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: Windows-95
: SOFTWARE: Patent in Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: FILING DATE: 08/26/94
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Van Amsterdam, John R.

```

```

: REGISTRATION NUMBER: 40,212
: REFERENCE/KEY NUMBER: 1045177005
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 617-720 3500
: TELEFAX: 617-720-2441
: INFORMATION FOR SEQ ID NO: 6:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 755 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: HYPOTHEetical: No
: ANTI-SENSE: NO
: ORIGINAL SOURCE: Homo sapiens
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 53..595
: US-08-791-495-6

Query Match          43.3%; Score 219; DB 1; Length 755;
Best Local Similarity 100.0%; Pred. No. 2.6e-50;
Matches 219; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 agatcccatgagcgtgagcgtctctgacagagatggaatgacccatgagagcccaaaag 60
|||||
DB 238 GGGTCGGATGGGATGAGTCTCTGGGAGATGGAAAGGTGTCCTGGGGGCCAGAGAG 297
|||||

QY 61 acacacacacacacacacacacacacacacacacacacacacacacacacacacacacac 120
|||||
DB 298 GCGGACACACACACACACACACACACACACACACACACACACACACACACACACACAC 457
|||||

QY 121 acacacacacacacacacacacacacacacacacacacacacacacacacacacacacac 180
|||||
DB 358 GCAAGGCGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 417
|||||

QY 181 tctaaacacacacacacacacacacacacacacacacacacacacacacacacacacacacac 219
|||||
DB 418 TCTGAAGACACTTACCGGCTGGGAGGACACACACACACACACACACACACACACACAC 456
|||||

```

```

RESULT 4
US-08-791-495-1
: Sequence 6; Application US/08791495
: Patent No. 5811519
: GENERAL INFORMATION:
: APPLICANT: Leth, Bernard
: APPLICANT: Lucas, Sophie
: APPLICANT: De Smet, Charles
: APPLICANT: Godelaine, Danielle
: APPLICANT: Boon-Fallour, Thierry
: TITLE OF INVENTION: LL-1 TUMOR SPECIFIC GENES
: NUMBER OF SEQUENCES: 14
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Wolf, Greenfield & Sacks, P.C.
: STREET: 600 Atlantic Avenue
: CITY: Boston
: STATE: MA
: COUNTRY: USA
: ZIP: 02210
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: Windows-95
: SOFTWARE: Patent in Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: FILING DATE: 08/26/94
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Van Amsterdam, John R.

```



```

1 Patent No. 5874564
2 GENERAL INFORMATION:
3 APPLICANT: Kim, Jungsuh P.
4 APPLICANT: Wares, John
5 APPLICANT: Young, Lawrence M.
6 APPLICANT: Park, Kirk E.
7 APPLICANT: Limon, Jeffrey M.
8 TITLE OF INVENTION: Hepatitis G Virus and Molecular
9 TITLE OF INVENTION: Hepatitis G Virus and Molecular
10 NUMBER OF SEQUENCES: 277
11 CORRESPONDENCE ADDRESS:
12 ADDRESS: Bellinger & Associates
13 STREET: 350 Cambridge Ave., Suite 250
14 CITY: Palo Alto
15 STATE: CA
16 COUNTRY: USA
17 ZIP: 94306
18 COMPUTER READABLE FORM:
19 MEDIUM TYPE: 3.5 inch disk
20 OPERATING SYSTEM: IBM PC compatible
21 SOFTWARE: Patent In Release #1.0, Version #1.25
22 CURRENT APPLICATION DATA:
23 APPLICATION NUMBER: 08/246,995
24 FILING DATE: 15 FEB 1995
25 CLASSIFICATION: 435
26 PRIOR APPLICATION DATA:
27 APPLICATION NUMBER: 08/246,995
28 FILING DATE: 15 FEB 1995
29 PRIOR APPLICATION DATA:
30 APPLICATION NUMBER: 08/246,995
31 FILING DATE: 15 FEB 1995
32 ATTORNEY/AGENT INFORMATION:
33 NAME: Fabian, Gary R.
34 REGISTRATION NUMBER: 44,875
35 REFERENCE/WORK NUMBER: 4600-0201-36/0100P11
36 TELEPHONE: (415) 324-0880
37 TELEFAX: (415) 324-0960
38 INFORMATION FOR SEQ ID NO: 182:
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000

```

Query Match: 7.0% Score: 65.24 ID: 2 Length: 9103
 Best Score: Similarity: 66.6% PctID: 1.4%

Software version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

em nucleotide nucleotide search, using sw model

Run on: May 1, 2001, 09:04:05 : Search time 3766.75 seconds

(without alignment)

1167.351 Million cell updates/sec

Hit list: US-09-341-829a-4_copy_250_755

Perfect scores: 506

Sequence: 1-nucleotide-sequences.....with gap-fixture code 506

Scoring table: HENRIY_NUP

Gapop 10.0 , Gapext 1.0

Sequences: 162517 seqs, 40064076 residues

Total number of hits satisfying chosen parameters: 1924704

Minimum hit seq length: 0

Maximum hit seq length: 20000000

Post processing: Minimum Match 0%

Maximum Match 100%

Listed first 45 summaries

Database 1

EST:

1: qb_08t11: *
2: qb_08t21: *
3: qb_08t41: *
4: qb_08t14: *
5: qb_08t51: *
6: qb_08t61: *
7: qb_08t71: *
8: qb_08t81: *
9: qb_08t91: *
10: qb_08t101: *
11: qb_08t111: *
12: qb_08t121: *
13: qb_08t131: *
14: qb_08t141: *
15: qb_08t151: *
16: qb_08t161: *
17: qb_08t171: *
18: qb_08t181: *
19: qb_08t191: *
20: qb_08t201: *
21: qb_08t211: *
22: qb_08t221: *
23: qb_08t231: *
24: qb_08t241: *
25: qb_08t251: *
26: qb_08t261: *
27: qb_08t271: *
28: qb_08t281: *
29: qb_08t291: *
30: qb_08t301: *
31: qb_08t311: *
32: qb_08t321: *
33: qb_08t331: *
34: qb_08t341: *
35: qb_08t351: *
36: qb_08t361: *
37: qb_08t371: *
38: qb_08t381: *
39: qb_08t391: *
40: qb_08t401: *
41: qb_08t411: *
42: qb_08t421: *
43: qb_08t431: *

44: em_08t101: *
45: em_08t111: *
46: em_08t121: *
47: em_08t131: *
48: em_08t141: *
49: em_08t151: *
50: em_08t161: *
51: em_08t171: *
52: em_08t181: *
53: em_08t191: *
54: em_08t201: *
55: em_08t211: *
56: em_08t221: *
57: em_08t231: *
58: em_08t241: *
59: em_08t251: *
60: em_08t261: *
61: em_08t271: *
62: em_08t281: *
63: em_08t291: *
64: em_08t301: *
65: em_08t311: *
66: em_08t321: *
67: em_08t331: *
68: em_08t341: *
69: em_08t351: *
70: em_08t361: *
71: em_08t371: *
72: em_08t381: *
73: em_08t391: *
74: em_08t401: *
75: em_08t411: *
76: em_08t421: *
77: em_08t431: *
78: em_08t441: *
79: em_08t451: *
80: em_08t461: *
81: em_08t471: *
82: em_08t481: *
83: em_08t491: *
84: em_08t501: *
85: em_08t511: *
86: em_08t521: *
87: em_08t531: *
88: em_08t541: *
89: em_08t551: *
90: em_08t561: *
91: em_08t571: *
92: em_08t581: *
93: em_08t591: *
94: em_08t601: *
95: em_08t611: *
96: em_08t621: *
97: em_08t631: *
98: em_08t641: *
99: em_08t651: *
100: em_08t661: *
101: em_08t671: *
102: em_08t681: *
103: em_08t691: *
104: em_08t701: *
105: em_08t711: *
106: em_08t721: *
107: em_08t731: *
108: em_08t741: *
109: em_08t751: *
110: em_08t761: *
111: em_08t771: *
112: em_08t781: *
113: em_08t791: *
114: em_08t801: *
115: em_08t811: *
116: em_08t821: *

117: qb_est48:*
 118: qb_est49:*
 119: qb_est50:*
 120: qb_est51:*
 121: qb_est52:*
 122: qb_est53:*
 123: qb_est54:*
 124: qb_est55:*
 125: qb_est56:*
 126: qb_est57:*
 127: qb_est58:*
 128: qb_est59:*
 129: qb_est60:*
 130: qb_est61:*
 131: qb_est62:*
 132: qb_est63:*
 133: qb_est64:*
 134: qb_est65:*
 135: qb_est66:*
 136: qb_est67:*
 137: qb_est68:*
 138: qb_est69:*
 139: qb_est70:*
 140: qb_est71:*
 141: qb_est72:*
 142: qb_est73:*
 143: qb_est74:*
 144: qb_est75:*
 145: qb_est76:*
 146: qb_est77:*
 147: qb_est78:*
 148: qb_est79:*
 149: qb_est80:*
 150: qb_est81:*
 151: qb_est82:*
 152: qb_est83:*
 153: qb_est84:*
 154: qb_est85:*
 155: qb_est86:*
 156: qb_est87:*
 157: qb_est88:*
 158: qb_est89:*
 159: qb_est90:*
 160: qb_est91:*
 161: qb_est92:*
 162: qb_est93:*
 163: qb_est94:*
 164: qb_est95:*
 165: qb_est96:*
 166: qb_est97:*
 167: qb_est98:*
 168: qb_est99:*
 169: qb_est100:*
 170: qb_est101:*
 171: qb_est102:*
 172: qb_est103:*
 173: qb_est104:*
 174: qb_est105:*
 175: qb_est106:*
 176: qb_est107:*
 177: qb_est108:*
 178: qb_est109:*
 179: qb_est110:*
 180: qb_est111:*
 181: qb_est112:*
 182: qb_est113:*
 183: qb_est114:*
 184: qb_est115:*
 185: qb_est116:*
 186: qb_est117:*
 187: qb_est118:*
 188: qb_est119:*
 189: qb_est120:*

em_qss_pln1:*
 em_qss_pln2:*
 em_qss_pro:*
 em_qss_fod1:*
 em_qss_fod2:*
 em_qss_fod3:*
 em_qss_fod4:*
 em_qss_fod5:*
 em_qss_vrt1:*
 em_qss_vrt2:*
 em_qss_vrt3:*
 qb_qss1:*
 qb_qss2:*
 qb_qss3:*
 qb_qss4:*
 qb_qss5:*
 qb_qss6:*
 qb_qss7:*
 qb_qss8:*
 qb_qss9:*
 qb_qss10:*
 qb_qss11:*
 qb_qss12:*
 qb_qss13:*
 qb_qss14:*
 qb_qss15:*
 qb_qss16:*
 qb_qss17:*
 qb_qss18:*
 qb_qss19:*
 qb_qss20:*
 qb_qss21:*
 qb_qss22:*
 qb_qss23:*
 qb_qss24:*
 qb_qss25:*
 qb_qss26:*
 qb_qss27:*
 qb_qss28:*
 qb_qss29:*
 qb_qss30:*
 qb_qss31:*
 qb_qss32:*
 qb_qss33:*
 qb_qss34:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed.
 and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	416.6	82.3	699	165	BE274123
2	345.2	68.2	596	166	BE398259
3	327.9	64.8	559	165	BE276194
4	313.2	61.9	749	166	BE391213
5	279.4	55.2	800	166	BE408892
6	225.2	44.5	674	166	BE388562
7	203.2	40.2	645	166	BE410697
8	185.2	36.6	852	166	BE385880
9	175.4	34.7	511	166	BE387918
10	156.8	31.0	483	166	BE410952
11	59.4	11.7	359	17	AL218223
12	56.4	11.1	534	31	AV663843
13	54	10.7	639	166	BE387002
14	51.6	10.2	925	229	CNS0091P
15	50.4	10.0	925	229	CNS0091P
16	49.6	9.8	498	111	AW161857
17	49.6	9.8	559	111	AW157381
18	49.6	9.8	714	31	AV703867

RESULT 12	AV66064.4	5.4 kb	mRNA	EST	25-AUG-2000
LOCUS	AV66064.4 Bos taurus brain telomerase RNA sequence				25-AUG-2000
DEFINITION	AV66064.4 Bos taurus brain telomerase RNA sequence				25-AUG-2000
ACCESSION	AV66064.4				
VERSION	AV66064.4				
KEYWORDS	EST				
SOURCE	EST				
ORGANISM	Bos taurus				
REFERENCE	1				
AUTHORS	Suzuki H, Hattori S, Takemura A, Hoshino A, Hoshino A				
TITLE	Brain telomerase RNA sequence				
JOURNAL	Published (2000)				
COMMENT	Brain telomerase RNA sequence				
FEATURES	Source				
ORIGIN	122 a 161 c 164 a 125 t				
QUALITY MATCH	11.1% Score 56.4; ID 41; Length 544				
BEST LOCAL SIMILARITY	57.4% Pred. No. 0.0001				
MATCHES	164; Conservative 0; Mismatches 86; Indels 5; Gaps 1				
QY	6				
QY	47				
QY	6				
QY	107				
QY	120				
QY	167				
QY	186				
QY	221				
RESULT 13	PE507002	6.5 kb	mRNA	EST	21-JUL-2000
LOCUS	PE507002.2 H. sapiens brain telomerase RNA sequence				21-JUL-2000
DEFINITION	PE507002.2 H. sapiens brain telomerase RNA sequence				21-JUL-2000
ACCESSION	PE507002.2				
VERSION	PE507002.2				
KEYWORDS	EST				
SOURCE	EST				
ORGANISM	Homo sapiens				
REFERENCE	1				
AUTHORS	Genoscope				
TITLE	Brain telomerase RNA sequence				
JOURNAL	Submitted (02-JUN-1999) Genoscope				
COMMENT	Brain telomerase RNA sequence				
FEATURES	Source				
ORIGIN	122 a 162 c 233 a 132 t				
QUALITY MATCH	13.7% Score 64; ID 166; Length 609				
BEST LOCAL SIMILARITY	90.8% Pred. No. 0.00046				
MATCHES	69; Conservative 0; Mismatches 5; Indels 2; Gaps 1				
QY	1				
QY	227				
QY	59				
QY	267				
RESULT 14	CNS009170	9.25 kb	mRNA	EST	03-JUN-1999
LOCUS	CNS009170.2 Drosophila melanogaster genomic survey sequence				03-JUN-1999
DEFINITION	CNS009170.2 Drosophila melanogaster genomic survey sequence				03-JUN-1999
ACCESSION	AL054013				
VERSION	AL054013				
KEYWORDS	EST				
SOURCE	EST				
ORGANISM	Drosophila melanogaster				
REFERENCE	1				
AUTHORS	Genoscope				
TITLE	Brain telomerase RNA sequence				
JOURNAL	Submitted (02-JUN-1999) Genoscope				
COMMENT	Brain telomerase RNA sequence				
FEATURES	Source				
ORIGIN	122 a 162 c 233 a 132 t				
QUALITY MATCH	13.7% Score 64; ID 166; Length 609				
BEST LOCAL SIMILARITY	90.8% Pred. No. 0.00046				
MATCHES	69; Conservative 0; Mismatches 5; Indels 2; Gaps 1				
QY	1				
QY	227				
QY	59				
QY	267				
RESULT 14	CNS009170	9.25 kb	mRNA	EST	03-JUN-1999
LOCUS	CNS009170.2 Drosophila melanogaster genomic survey sequence				03-JUN-1999
DEFINITION	CNS009170.2 Drosophila melanogaster genomic survey sequence				03-JUN-1999
ACCESSION	AL054013				
VERSION	AL054013				
KEYWORDS	EST				
SOURCE	EST				
ORGANISM	Drosophila melanogaster				
REFERENCE	1				
AUTHORS	Genoscope				
TITLE	Brain telomerase RNA sequence				
JOURNAL	Submitted (02-JUN-1999) Genoscope				
COMMENT	Brain telomerase RNA sequence				
FEATURES	Source				
ORIGIN	122 a 162 c 233 a 132 t				
QUALITY MATCH	13.7% Score 64; ID 166; Length 609				
BEST LOCAL SIMILARITY	90.8% Pred. No. 0.00046				
MATCHES	69; Conservative 0; Mismatches 5; Indels 2; Gaps 1				
QY	1				
QY	227				
QY	59				
QY	267				
RESULT 14	CNS009170	9.25 kb	mRNA	EST	03-JUN-1999
LOCUS	CNS009170.2 Drosophila melanogaster genomic survey sequence				03-JUN-1999
DEFINITION	CNS009170.2 Drosophila melanogaster genomic survey sequence				03-JUN-1999
ACCESSION	AL054013				
VERSION	AL054013				
KEYWORDS	EST				
SOURCE	EST				
ORGANISM	Drosophila melanogaster				
REFERENCE	1				
AUTHORS	Genoscope				
TITLE	Brain telomerase RNA sequence				
JOURNAL	Submitted (02-JUN-1999) Genoscope				
COMMENT	Brain telomerase RNA sequence				
FEATURES	Source				
ORIGIN	122 a 162 c 233 a 132 t				
QUALITY MATCH	13.7% Score 64; ID 166; Length 609				
BEST LOCAL SIMILARITY	90.8% Pred. No. 0.00046				
MATCHES	69; Conservative 0; Mismatches 5; Indels 2; Gaps 1				
QY	1				
QY	227				
QY	59				
QY	267				
RESULT 14	CNS009170	9.25 kb	mRNA	EST	03-JUN-1999
LOCUS	CNS009170.2 Drosophila melanogaster genomic survey sequence				03-JUN-1999
DEFINITION	CNS009170.2 Drosophila melanogaster genomic survey sequence				03-JUN-1999
ACCESSION	AL054013				
VERSION	AL054013				
KEYWORDS	EST				
SOURCE	EST				
ORGANISM	Drosophila melanogaster				
REFERENCE	1				
AUTHORS	Genoscope				
TITLE	Brain telomerase RNA sequence				
JOURNAL	Submitted (02-JUN-1999) Genoscope				
COMMENT	Brain telomerase RNA sequence				
FEATURES	Source				
ORIGIN	122 a 162 c 233 a 132 t				
QUALITY MATCH	13.7% Score 64; ID 166; Length 609				
BEST LOCAL SIMILARITY	90.8% Pred. No. 0.00046				
MATCHES	69; Conservative 0; Mismatches 5; Indels 2; Gaps 1				
QY	1				
QY	227				
QY	59				
QY	267				

collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see <http://www.fruitfly.org/BDGP/melanogaster>. The BDGP melanogaster BAC library was prepared by Kazutoyo Asanuma and Aaron Margosier in Peter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCL 98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's pl and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at <http://www.bacpac.org/bacpac/bacpac.html>.

FEATURES
Source
Location/Qualifiers
1..925
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone_lib="RPCL 98"
/clone="BAC19016"
/note="end : TET3"
BASE COUNT 120 a 61 c 61 g 172 t 511 others
ORIGIN

Query Match 10.9% Score 51.6; DB 229; Length 925;
Best Local Similarity 14.9%; Pred. No. 0.0021;
Matches 41; Conservative 14; Mismatches 149; Indels 9; Gaps 0;
QY 85 aagatcagatgacatcctctcagccacatggaagggagatgagcagatcctcagc 144
T
DQ 918 aag 969
T
QY 145 caaatatgac 974
T
DQ 878 aac 799
T
QY 205 caac 984
T
DQ 798 aac 799
T
QY 205 aatgatttgaatgagatgagatgagatgagatgagatgagatgagatgag 924
T
DQ 734 aag 679
T
QY 425 aac 984
T
DQ 674 aac 679
T
QY 485 aagcagac 414
T
DQ 618 aac 589
T T

RESULT 15
CNS0041P 925 bp DNA GSS 03 JUN 1999
Locus Drosophila melanogaster genome survey sequence TET3 end of BAC #
BAC19016 of RPCL-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION AL054013.1 GI:4934461
VERSION AL054013.1
KEYWORDS GSS.
SOURCE fruit fly.
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 925)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (02 JUN 1999) Genoscope, Centre National de Recherche
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)

- Web : www.genoscope.cns.fr
Sequencing of this BAC end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see <http://www.fruitfly.org/BDGP/melanogaster>. The BDGP melanogaster BAC library was prepared by Kazutoyo Asanuma and Aaron Margosier in Peter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCL 98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's pl and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at <http://www.bacpac.org/bacpac/bacpac.html>.

FEATURES
Source
Location/Qualifiers
1..925
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone_lib="RPCL 98"
/clone="BAC19016"
/note="end : TET3"
BASE COUNT 120 a 61 c 61 g 172 t 511 others
ORIGIN

Query Match 10.9% Score 50.4; DB 229; Length 925;
Best Local Similarity 11.7%; Pred. No. 0.0044;
Matches 35; Conservative 149; Mismatches 116; Indels 0; Gaps 0;
QY 1 aagatcagatgacatcctctcagccacatggaagggagatgagcagatcctcagc 60
T
DQ 884 aag 662
T
QY 61 aagcagac 120
T
DQ 733 aag 722
T
QY 121 aagcagac 180
T
DQ 723 aag 782
T
QY 193 aagcagac 240
T
DQ 784 aag 842
T
QY 341 aagcagac 400
T
DQ 843 aag 902
T T

Search completed: May 1, 2001, 04:04:08
Job time: 11406 sec

GenCore version 4.5
Copyright (c) 1993-2000 Compugen Ltd.

on nucleotide nucleotide search, using sw model

Run on: May 1, 2001, 08:51:24 ; Search time: 2512.89 Seconds
(without alignments)
123,247 Million cell updates/sec

File: us_09_341_829a_4_copy_488_508
Perfect score: 21
Sequence: 1-annanur-qetannuqqatq_21

Scoring table: IDENTITY, 80%
Gapop 10,0 ; Gapext 1,0

Searched: 120820 seqs, 787029662 residues
Total number of hits satisfying chosen parameters: 2566470

Minimum hit seq length: 0

Maximum hit seq length: 20000000

Best processing: Maximum Match on
CDS and first 45 summary lines

Database:

1: all_baz1*
2: all_baz2*
3: all_baz3*
4: all_baz4*
5: all_baz5*
6: all_baz6*
7: all_baz7*
8: all_baz8*
9: all_baz9*
10: all_baz10*
11: all_baz11*
12: all_baz12*
13: all_baz13*
14: all_baz14*
15: all_baz15*
16: all_baz16*
17: all_baz17*
18: all_baz18*
19: all_baz19*
20: all_baz20*
21: all_baz21*
22: all_baz22*
23: all_baz23*
24: all_baz24*
25: all_baz25*
26: all_baz26*
27: all_baz27*
28: all_baz28*
29: all_baz29*
30: all_baz30*
31: all_baz31*
32: all_baz32*
33: all_baz33*
34: all_baz34*
35: all_baz35*
36: all_baz36*
37: all_baz37*
38: all_baz38*
39: all_baz39*
40: all_baz40*
41: all_baz41*
42: all_baz42*
43: all_baz43*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARY

Result No.	Score	Match	Length	bits	Prod. No.	Description
1	21	100.0	994	9	AX024689	AX024689 Sequence
2	21	100.0	994	92	BSA012845	BSA012845 Homo sapi
3	21	100.0	1002	9	AK042535	AK042535 Sequence
4	21	100.0	1002	92	BSA224040	BSA224040 Homo sapi
5	21	100.0	1002	91	BSA224094	BSA224094 Homo sapi
6	21	100.0	112874	78	AF277415	AF277415 Homo sapi
7	17.8	84.8	2941	94	HS0110694	HS0110694 Human, MA22-
8	17.8	84.8	40438	88	AC087410	AC087410 Homo sapi
9	17.8	84.8	48927	93	BS009568	BS009568 Human, Xp28
10	17.8	84.8	73360	93	BS006083	BS006083 Human, cont1
11	17.8	84.8	111560	87	AC016940	AC016940 Homo sapi

12	17.8	84.8	118440	87	AC016939	Homo sapi
13	17.8	84.8	214894	70	AC025751	Mus muscu
14	17.4	82.9	4900	14	PC01PAM6	X54257 P. chrysosp
15	17.4	82.9	10522	1	AE034514	AL030454 Pseudomon
16	17.4	82.9	339681	2	AF003009	AF003009 Mesorhizo
17	17.4	82.9	345783	2	AF003001	AF003001 Mesorhizo
18	16.8	80.0	58280	4	MT0014	AL021646 Mycobacte
19	16.8	80.0	118947	60	AC006321	AC006321 Homo sapi
20	16.8	80.0	124040	83	CNS01081	AL121808 Homo sapi
21	16.8	80.0	124337	92	HSR01609	AL121825 Human DNA
22	16.8	80.0	132698	60	AC005051	AC005051 Homo sapi
23	16.8	80.0	167457	72	AC037478	AC037478 Homo sapi
24	16.8	80.0	150744	85	AC005038	AC005038 Homo sapi
25	16.8	80.0	190846	86	AC0056372	AC0056372 Homo sapi
26	16.8	80.0	193038	61	AC008742	AC008742 Homo sapi
27	16.8	80.0	213355	63	AC012451	AC012451 Homo sapi
28	16.4	78.1	964	94	ME01278	AC074111 Mouse mRNA
29	16.4	78.1	984	94	ME01278	AC074111 Mouse mRNA
30	16.4	78.1	1464	94	AF093403	AF093403 Mus muscu
31	16.4	78.1	2436	94	AF288694	AF288694 Mus muscu
32	16.4	78.1	3528	88	AF014643	AF014643 Homo sapi
33	16.4	78.1	4538	88	AF147709	AF147709 Homo sapi
34	16.4	78.1	23610	62	AC011805	AC011805 Homo sapi
35	16.4	78.1	41604	93	HS04119	AC011932 Homo sapi
36	16.4	78.1	126138	92	HS04119	AC011932 Homo sapi
37	16.4	78.1	134036	86	AC007845	Y18090 Homo sapien
38	16.4	78.1	136030	90	AL350815	AL007845 Homo sapi
39	16.4	78.1	152137	89	AL357252	AL359815 Homo sapi
40	16.4	78.1	157023	82	AL513129	AL357252 Homo sapi
41	16.4	78.1	171878	85	AC004882	AL513129 Homo sapi
42	16.4	78.1	180187	68	AC024093	AC004882 Homo sapi
43	16.4	78.1	183920	80	AL459510	AC023693 Homo sapi
44	16.4	78.1	194982	81	AL451069	AL459510 Homo sapi
45	16.4	78.1	201377	75	AC073767	AL451069 Homo sapi
46	16.4	78.1	201377	75	AC073767	AC073767 Mus muscu
47	16.4	78.1	201377	75	AC073767	AC073767 Mus muscu
48	16.4	78.1	201377	75	AC073767	AC073767 Mus muscu
49	16.4	78.1	201377	75	AC073767	AC073767 Mus muscu
50	16.4	78.1	201377	75	AC073767	AC073767 Mus muscu
51	16.4	78.1	201377	75	AC073767	AC073767 Mus muscu
52	16.4	78.1	201377	75	AC073767	AC073767 Mus muscu
53	16.4	78.1	201377	75	AC073767	AC073767 Mus muscu
54	16.4	78.1	201377	75	AC073767	AC073767 Mus muscu
55	16.4	78.1	201377	75	AC073767	AC073767 Mus muscu
56	16.4	78.1	201377	75	AC073767	AC073767 Mus muscu
57	16.4	78.1	201377	75	AC073767	AC073767 Mus muscu
58	16.4	78.1	201377	75	AC073767	AC073767 Mus muscu
59	16.4	78.1	201377	75	AC073767	AC073767 Mus muscu
60	16.4	78.1	201377	75	AC073767	AC073767 Mus muscu
61	16.4	78.1	201377	75	AC073767	AC073767 Mus muscu
62	16.4	78.1	201377	75	AC073767	AC073767 Mus muscu
63	16.4	78.1	201377	75	AC073767	AC073767 Mus muscu
64	16.4	78.1	201377	75	AC073767	AC073767 Mus muscu
65	16.4	78.1	201377	75	AC073767	AC073767 Mus muscu
66	16.4	78.1	201377	75	AC073767	AC073767 Mus muscu
67	16.4	78.1	201377	75	AC073767	AC073767 Mus muscu
68	16.4	78.1	201377	75	AC073767	AC073767 Mus muscu
69	16.4	78.1	201377	75	AC073767	AC073767 Mus muscu
70	16.4	78.1	201377	75	AC073767	AC073767 Mus muscu
71	16.4	78.1	201377	75	AC073767	AC073767 Mus muscu
72	16.4	78.1	201377	75	AC073767	AC073767 Mus muscu
73	16.4	78.1	201377	75	AC073767	AC073767 Mus muscu
74	16.4	78.1	201377	75	AC073767	AC073767 Mus muscu
75	16.4	78.1	201377	75	AC073767	AC073767 Mus muscu
76	16.4	78.1	201377	75	AC073767	AC073767 Mus muscu
77	16.4	78.1	201377	75	AC073767	AC073767 Mus muscu
78	16.4	78.1	201377	75	AC073767	AC073767 Mus muscu
79	16.4	78.1	201377	75	AC073767	AC073767 Mus muscu
80	16.4	78.1	201377	75	AC073767	AC073767 Mus muscu
81	16.4	78.1	201377	75	AC073767	AC073767 Mus muscu
82	16.4	78.1	201377	75	AC073767	AC073767 Mus muscu
83	16.4	78.1	201377	75	AC073767	AC073767 Mus muscu
84	16.4	78.1	201377	75	AC073767	AC073767 Mus muscu
85	16.4	78.1	201377	75	AC073767	AC073767 Mus muscu
86	16.4	78.1	201377	75	AC073767	AC073767 Mus muscu
87	16.4	78.1	201377	75	AC073767	AC073767 Mus muscu
88	16.4	78.1	201377	75	AC073767	AC073767 Mus muscu
89	16.4	78.1	201377	75	AC073767	AC073767 Mus muscu
90	16.4	78.1	201377	75	AC073767	AC073767 Mus muscu
91	16.4	78.1	201377	75	AC073767	AC073767 Mus muscu
92	16.4	78.1	201377	75	AC073767	AC073767 Mus muscu
93	16.4	78.1	201377	75	AC073767	AC073767 Mus muscu
94	16.4	78.1	201377	75	AC073767	AC073767 Mus muscu
95	16.4	78.1	201377	75	AC073767	AC073767 Mus muscu
96	16.4	78.1	201377	75	AC073767	AC073767 Mus muscu
97	16.4	78.1	201377	75	AC073767	AC073767 Mus muscu
98	16.4	78.1	201377	75	AC073767	AC073767 Mus muscu
99	16.4	78.1	201377	75	AC073767	AC073767 Mus muscu
100	16.4	78.1	201377	75	AC073767	AC073767 Mus muscu

[illegible]

Garrelli, J. H., Correll, L. L., Giovanna, W., Harris, K., Hernandez, J., Hodgson, A., Holmes, M., Holloway, C., Hosak, H., Jackson, L. B., Jackson, L., Jia, Y., Jones, M., Kelly, S., Kondejowski, N., Koud, Y., Kovari, C., Leal, B., Li, Z., Lichtarge, O., Liu, J., Liu, W., Logan, O., Lozano, R. J., Lu, J., Lurie, R., Martin, R., Martinez, C., McLeod, M. P., Mei, G., Moran, M., Morris, S., Rast, J., Wilson, A., Wang, S., Naylor, N., Nguyen, S., Nuss, G., Pamrell, L. K., Parfisch, A., Paxton, S., Payne, B., Perez, L., Pichler, L., Quiles, M., Rashid, N. D., Reiter, D., Rives, M., Samuel, S., Say, J., Scherer, S., Shah, E., Shen, H., Simon, M., Sparks, A., Stamps, A., Surugan, R., Taper, P., Taylor, T., Vasquez, I., Vinson, P., Vo, Q., Wahbeh, M., Watlington, S., Weinstein, G., Weinstein, I. P., Williams, A., Worley, E., Wren, J., Wrenston, G., Yu, W., Zhou, X., Nelson, D., and Gibbs, R.

Direct Submission
Unpublished
2 (bases 1 to 111560)
Worley, K.C.

Direct Submission
Submitted (09-DEC-1999) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
4 (bases 1 to 111560)
Worley, K.C.

Direct Submission
Submitted (01-APR-2000) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
on Apr 1, 2000 this sequence version replaced gi:7007873.
INFO: http://www.bgi.com/seq/ or email
us-help@bcm.tmc.edu

CLONE LENGTH: This sequence does not necessarily represent the entire insert of this clone. Overlapping regions of clones are only sequenced and submitted once, so the sequence for the remainder of the insert may be found in the record for the adjacent clones. Overlapping clones are noted at the beginning and end of the features listing.

ANNOTATION OF FEATURES:

STES are identified using STES (Shen, 1997, 550), searched a local database that includes entries from dbSTS, GDB, and local mapping efforts.

Repeats are identified using RepeatMasker (A. Smit and P. Green, unpublished) for human and mouse sequences.

Genes and region of sequence similarity are identified by BLAST (Altschul, et al., 1990, 403) similarity (expectation 1e-35) to the EST and cDNA sequences. Genes are annotated at least two exons flanked by consensus splice sites that maintained sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 1 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found at <http://www.bgi.com/seq/quality.htm>.

QUALSTAT-REPORT-----

----- Summary Statistics -----
Contig length: 111560
Phrap values in estimate: 111560
Average error rate (RCM Phrap estimate): 0.000108566
Fraction of Phrap values less than 40 : 0.0136415

Number of consensus changed edits:
Number of N's in consensus :

2
0

----- Consensus changing edits -----

Feature: Consensus
35684 aaaaaaaaaa(n)taaacaaaaa aaaaaaaaaa(a)taaacaaaaa
82852 aaaaaaaaaa(n)taaacaaaaa aaaaaaaaaa(a)taaacaaaaa

----- Distribution of Quality < 40 bases -----

# bases	1000	900	800	700	600	500	400	300	200	100	0	5	10	15	20	25	30	35	40
1000																			
900																			
800																			
700																			
600																			
500																			
400																			
300																			
200																			
100																			
0																			

Version: 1.01 qxf0.

FEATURES

Location/Qualifiers
1..111560

Source: "Homo sapiens"

gb|AF012420.1|taxon:"Homo sapiens"

gb|AF012420.1|taxon:"Homo sapiens"

gb|AF012420.1|taxon:"Homo sapiens"

gb|AF012420.1|taxon:"Homo sapiens"

gb|AF012420.1|taxon:"Homo sapiens"

gb|AF012420.1|taxon:"Homo sapiens"

gb|AF012420.1|taxon:"Homo sapiens"

gb|AF012420.1|taxon:"Homo sapiens"

gb|AF012420.1|taxon:"Homo sapiens"

gb|AF012420.1|taxon:"Homo sapiens"

gb|AF012420.1|taxon:"Homo sapiens"

gb|AF012420.1|taxon:"Homo sapiens"

gb|AF012420.1|taxon:"Homo sapiens"

gb|AF012420.1|taxon:"Homo sapiens"

gb|AF012420.1|taxon:"Homo sapiens"

gb|AF012420.1|taxon:"Homo sapiens"

gb|AF012420.1|taxon:"Homo sapiens"

gb|AF012420.1|taxon:"Homo sapiens"

gb|AF012420.1|taxon:"Homo sapiens"

gb|AF012420.1|taxon:"Homo sapiens"

gb|AF012420.1|taxon:"Homo sapiens"

gb|AF012420.1|taxon:"Homo sapiens"

gb|AF012420.1|taxon:"Homo sapiens"

gb|AF012420.1|taxon:"Homo sapiens"

gb|AF012420.1|taxon:"Homo sapiens"

gb|AF012420.1|taxon:"Homo sapiens"

gb|AF012420.1|taxon:"Homo sapiens"

gb|AF012420.1|taxon:"Homo sapiens"

gb|AF012420.1|taxon:"Homo sapiens"

gb|AF012420.1|taxon:"Homo sapiens"

gb|AF012420.1|taxon:"Homo sapiens"

gb|AF012420.1|taxon:"Homo sapiens"

gb|AF012420.1|taxon:"Homo sapiens"

gb|AF012420.1|taxon:"Homo sapiens"

gb|AF012420.1|taxon:"Homo sapiens"

gb|AF012420.1|taxon:"Homo sapiens"

gb|AF012420.1|taxon:"Homo sapiens"

gb|AF012420.1|taxon:"Homo sapiens"

gb|AF012420.1|taxon:"Homo sapiens"

gb|AF012420.1|taxon:"Homo sapiens"

gb|AF012420.1|taxon:"Homo sapiens"

gb|AF012420.1|taxon:"Homo sapiens"

gb|AF012420.1|taxon:"Homo sapiens"

gb|AF012420.1|taxon:"Homo sapiens"

gb|AF012420.1|taxon:"Homo sapiens"

gb|AF012420.1|taxon:"Homo sapiens"

gb|AF012420.1|taxon:"Homo sapiens"

gb|AF012420.1|taxon:"Homo sapiens"

Search completed: May 1, 2001, 03:51:35
Job time: 11628 sec

PA (BEH) BOEHRINGER INGELHEIM INT GMBH.
 XX (OHE-) UNIV HOSPITAL LEIDEN.
 PI Schrier PL, Aaronsse CA, Heider K, Klade C;
 XX WPI: 2000-039685/29.
 DR P-PSDB: Y70861.
 XX
 XX Tumor associated antigen useful for cancer immunotherapy is encoded by
 PI the open reading frame of LAGE-1 (a tumor-specific antigen) cDNA -
 XX
 PS Example 1: Page 58-59; 74pp; English.
 XX
 XX The present sequence is the cDNA encoding human LAGE-1 unspliced variant,
 CC LAGE-1L, a tumor-specific antigen. It decodes the LAGE-1L protein from
 CC open reading frame-3 (ORF) and the CAMPE protein (cytotoxic T lymphocyte
 CC cell) recognized antigen (cancer-associated protein 1, the LAGE-1L) protein
 CC from ORF 1 is identical to the CAMPE protein. LAGE-1 is not expressed in
 CC healthy tissues except in testis and placenta. It also shows homology
 CC with the NY-ESO-1, a melanoma specific tumor antigen. The tumor-
 CC associated antigen displayed on melanoma cells is recognized by cytotoxic
 CC T lymphocytes. This sequence has anticancer activity. CAMPE tumor
 CC antigen and immunogenic peptides derived from it are useful for cancer
 CC immunotherapy. They have the potential to induce an immune response, by
 CC eliciting a CTL response. The DNA molecule is used for the construction
 CC of recombinant or fusion proteins.
 XX
 SQ Sequence 993 BP; 184 A; 295 C; 444 G; 180 T; 0 other;

Query Match 100.0%; Score 21; DB 21; Length 993;
 Best Local Similarity 100.0%; Pred. No. 1,2;
 Matches 21: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 aaaaaagcgcctqqcqaatg 21
 IIIIIIIIIIIIIIIIIIII
 bb 47% aaaaaagcgcctqqcqaatg 499

RESL1 2
 V0346
 ID V0346 standard: cDNA; 1002 BP.
 XX
 AC V0346;
 XX
 DI 27-OCT-1998 (first entry)
 XX
 XX Human LAGE-1 clone 2 nucleotide sequence.
 DE
 XX Human LAGE-1: tumor associated protein; LL 1, diagnosis, tumour ds.
 KW Homo sapiens.
 OS Key Location/Qualifiers
 FH CDS 65..667
 FI /start a
 FI /product "LAGE-1"
 XX
 IN W09832855.A1.
 XX
 PU 30-JUL-1998.
 XX
 XX 27 JAN-1998; 98WO-0501445.
 XX
 XX 27-JAN-1997; 970S-0791495.
 XX
 XX (LAGE-1) L00WIG INSI CANCER RES.
 XX
 PI Beau-Falloux T, De Smet C, Godetaine D, Letho B;
 PI Lucas S;
 XX
 XX WPI: 1998-427951/30.
 DR P-PSDB: W6663.

XX New isolated LAGE-1 tumour associated nucleic acids used to
 PI develop products for the diagnosis and treatment of LAGE-1
 PI associated disorders, particularly tumours
 XX
 PS Claim 1: Page 50-52; 73pp; English.
 XX
 XX The present sequence encodes LAGE-1 tumour associated protein (LAP).
 CC The present invention also describes: (1) a method for treating a
 CC subject with a disorder characterised by expression of a LAGE-1 nucleic
 CC acid molecule or an expression product, comprising administering a
 CC subject autologous cytolytic T cells to ameliorate the disorder, where
 CC the cytolytic T cells are specific for epitopes of an HLA molecule and
 CC a LAGE-1 TAP or an immunogenic fragment; (2) a method for treating a
 CC subject with a disorder characterised by expression of a LAGE-1 nucleic
 CC acid molecule or an expression product, comprising administering a
 CC LAGE-1 TAP or an immunogenic fragment to ameliorate the disorder; and
 CC (3) a method for selectively enriching a population of T cells with
 CC cytolytic T cells specific for a LAGE-1 TAP comprising contacting an
 CC isolated population of T cells with an agent presenting a complex of a
 CC LAGE-1 TAP or an immunogenic fragment and a HLA presenting molecule to
 CC selectively enrich the isolated population of T cells with the cytolytic
 CC T cells. The methods and products from the present invention can be used
 CC for the diagnosis and treatment of LAGE-1 associated disorders.
 XX
 SQ Sequence 1002 BP; 179 A; 302 C; 436 G; 185 T; 0 other;

Query Match 100.0%; Score 21; DB 19; Length 1002;
 Best Local Similarity 100.0%; Pred. No. 1,2;
 Matches 21: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 aaaaaagcgcctqqcqaatg 21
 IIIIIIIIIIIIIIIIIIII
 bb 48% aaaaaagcgcctqqcqaatg 508

RESL1 3
 C68127
 ID C68127 standard: cDNA; 1769 BP.
 XX
 AC C68127;
 XX
 DI 20 FEB-2001 (first entry)
 XX
 XX Human secreted protein cDNA sequence #47.
 DE
 XX Cytostatic; immunosuppressive; neoplastic; neuroprotective; antiviral;
 KW antidiabetic; hepatotropic; antidiabetic; antineoplastic; antineuro;
 KW pulmonary; antineoplastic; antibacterial; antifungal; antiparasitic;
 KW cardiac; gene therapy; cancer; immune disorder; cardiovascular disorder;
 KW neurological disease; infection; human; secreted protein; ss.
 XX
 OS Homo sapiens.
 XX
 IN W1200058335.A1.
 XX
 PU 05-OCT-2000.
 XX
 XX 22 MAR-2000; 2000WO-0507534.
 XX
 XX 26 MAR-1999; 990S-0126598.
 XX
 XX 22 DEC-1999; 990S-0171504.
 XX
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX (R08K/) ROSEN C A.
 XX
 PI Rosen CA, Ruben SM, Komatsu S G;
 PI WPI: 2000-611702/58.
 DR P-PSDB: B37394.
 XX


```

F1 stem_loop      /*tag d
F1 120..568
F1 /*tag e
F1 569..1665
F1 /*tag c
F1 misc_structure 120..124
F1 /*tag d
F1 /note= "bases potentially involved in hairpin
F1 formation"
F1 misc_structure 140..134
F1 /*tag c
F1 /note= "bases potentially involved in hairpin
F1 formation"
F1 misc_structure 574..578
F1 /*tag f
F1 /note= "bases potentially involved in hairpin
F1 formation"
F1 misc_structure 584..588
F1 /*tag d
F1 /note= "bases potentially involved in hairpin
F1 formation"
F1 polyA_signal    1647..1642
F1 /*tag h
F1
F1 US9506126-A.
F1
F1 09-APR-1996.
F1
F1 25-FEB-1988; 8805-0160416.
F1
F1 01-DEC-1992; 9203-0594647.
F1 25-FEB-1988; 8805-0160416.
F1 14-JUL-1989; 8905-0379076.
F1 14-JUL-1990; 9005-0554759.
F1 18-DEC-1992; 9203-0139273.
F1
F1 (GENE) GEN HOSPITAL CORP.
F1
F1 Arutfo A. Seed B;
F1
F1 833..3596 20..275/20.
F1 P-PSDB; R91444.
F1
F1 Cloning of cDNA encoding cell surface antigen - useful for isolation
F1 of diagnostic and therapeutic proteins
F1
F1 Example 4; Fig 8A-8B; 79pp; English.
F1
F1 A cDNA clone (T14708) codes for human CD7 (R91434), a cell
F1 surface antigen associated with T-lymphocytes whose physiological
F1 role is not yet known. It was obtained by constructing an expression
F1 library in COS cells using cDNA derived from human T-cell tumour line
F1 HOP-ALL and vector pB3 (see also T14703), and screening of the library
F1 using antibody-coated plates. This immunoselection cloning method,
F1 developed to clone genes for cell surface antigens of human
F1 lymphocytes (see also T14703 04 and T14706 26), has general appli.
F1 Cell surface antigens are obtain. for diagnostic and therapeutic use.
F1
F1 Sequence 1665 BP; 402 A; 614 C; 479 G; 270 T; 0 other;

```

```

Query Match      75.2% Score 15.8; DB 17; Length 1665;
Best Local Similarity 89.5%; Prob. No. 2.1e-02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 1 aaaaaaagcgcctgcgcga 19
   11111111111111111111
DB 1260 AGGAAGCTGCTGCGCGGA 1242

```

```

RESULT 10
114708/c
ID 114708 standard; cDNA; 1665 BP.
AC 114708;
XX
XX to cct-1996 (first entry)
XX
XX Human CD7 cDNA.
XX
XX Cell surface antigen; cloning; immunoselection; immunotherapy;
XX therapy; diagnosis; vector; pB3, CD7, CD8, T lymphocyte, ss.
XX Homo sapiens.
XX
XX Key Location/Qualifiers
F1 exon 1..140
F1 /*tag a
F1 /exon_start 58..60

```

```

RESULT 11
V63446/c
ID V63446 standard; cDNA; 1665 BP.
XX

```


117: qb_est48:*
 118: qb_est49:*
 119: qb_est50:*
 120: qb_est51:*
 121: qb_est52:*
 122: qb_est53:*
 123: qb_est54:*
 124: qb_est55:*
 125: qb_est56:*
 126: qb_est57:*
 127: qb_est58:*
 128: qb_est59:*
 129: qb_est60:*
 130: qb_est61:*
 131: qb_est62:*
 132: qb_est63:*
 133: qb_est64:*
 134: qb_est65:*
 135: qb_est66:*
 136: qb_est67:*
 137: qb_est68:*
 138: qb_est69:*
 139: qb_est70:*
 140: qb_est71:*
 141: qb_est72:*
 142: qb_est73:*
 143: qb_est74:*
 144: qb_est75:*
 145: qb_est76:*
 146: qb_est77:*
 147: qb_est78:*
 148: qb_est79:*
 149: qb_est80:*
 150: qb_est81:*
 151: qb_est82:*
 152: qb_est83:*
 153: qb_est84:*
 154: qb_est85:*
 155: qb_est86:*
 156: qb_est87:*
 157: qb_est88:*
 158: qb_est89:*
 159: qb_est90:*
 160: qb_est91:*
 161: qb_est92:*
 162: qb_est93:*
 163: qb_est94:*
 164: qb_est95:*
 165: qb_est96:*
 166: qb_est97:*
 167: qb_est98:*
 168: qb_est99:*
 169: qb_est100:*
 170: qb_est101:*
 171: qb_est102:*
 172: qb_est103:*
 173: qb_est104:*
 174: qb_est105:*
 175: qb_est106:*
 176: qb_est107:*
 177: qb_est108:*
 178: qb_est109:*
 179: qb_est110:*
 180: qb_est111:*
 181: qb_est112:*
 182: qb_est113:*
 183: qb_est114:*
 184: qb_est115:*
 185: qb_est116:*
 186: qb_est117:*
 187: qb_est118:*
 188: qb_est119:*
 189: qb_est120:*

190: em_qss_pln1:*
 191: em_qss_pln2:*
 192: em_qss_pro:*
 193: em_qss_rnd1:*
 194: em_qss_rnd2:*
 195: em_qss_rnd3:*
 196: em_qss_rnd4:*
 197: em_qss_rnd5:*
 198: em_qss_vrt1:*
 199: em_qss_vrt2:*
 200: em_qss_vrt3:*
 201: qb_qss1:*
 202: qb_qss2:*
 203: qb_qss3:*
 204: qb_qss4:*
 205: qb_qss5:*
 206: qb_qss6:*
 207: qb_qss7:*
 208: qb_qss8:*
 209: qb_qss9:*
 210: qb_qss10:*
 211: qb_qss11:*
 212: qb_qss12:*
 213: qb_qss13:*
 214: qb_qss14:*
 215: qb_qss15:*
 216: qb_qss16:*
 217: qb_qss17:*
 218: qb_qss18:*
 219: qb_qss19:*
 220: qb_qss20:*
 221: qb_qss21:*
 222: qb_qss22:*
 223: qb_qss23:*
 224: qb_qss24:*
 225: qb_qss25:*
 226: qb_qss26:*
 227: qb_qss27:*
 228: qb_qss28:*
 229: qb_qss29:*
 230: qb_qss30:*
 231: qb_qss31:*
 232: qb_qss32:*
 233: qb_qss33:*
 234: qb_qss34:*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query %		DB	ID	Description
		Match	Length			
1	21	100.0	559	165	HE276194	HE276194 60124452
2	21	100.0	585	166	HE390259	HE390259 601245141
3	21	100.0	749	166	HE391232	HE391232 601245220
4	18.4	87.6	555	165	HE337694	HE337694 804047099
5	18.4	87.6	569	165	HE337695	HE337695 804047099
6	18	85.7	242	115	AW436749	AW436749 77205 MAR
7	18	85.7	242	115	AW436782	AW436782 77257 MAR
8	18	85.7	465	164	HE233043	HE233043 109049 MA
9	18	85.7	508	168	HE710007	HE710007 M1-0-AY1-
10	18	85.7	511	166	HE387918	HE387918 601242144
11	18	85.7	817	241	CNS03K53	AL247741 Tetraodon
12	18	85.7	964	231	CNS03X16	AL265057 Tetraodon
13	17.4	82.9	429	158	HE4362	HE4362 mdh9008.r1
14	17.4	82.9	674	166	HE388562	HE388562 601241855
15	17	81.0	951	232	CNS0512K	AL349545 Tetraodon
16	17	81.0	1101	229	CNS00198	AL068027 Drosophila
17	15.8	99.0	66	249	CNS02DHU	AL142671 Tetraodon
18	16.8	80.0	255	160	HE575394	HE575394 80575394

REFERENCE
AUTHORS
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus
1 (bases 1 to 465)
Fahrenkrug, S.C., Freking, B.A., Rohrer, G.A., Smith, T.P.L., Casas, E.,
Stone, K.L., Beaton, M.P., Grosse, W.M., Bennett, G.A., Laegreid, W.W.
and Kvelo, J.W.
TITLE
Design and use of two pooled tissue normalized cDNA libraries for
EST discovery in swine
JOURNAL
Unpublished (2000)
COMMENT
Contact: Smith, IPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smithhermail@mar.usda.gov
Single pass sequencing. Bases called and all trimmed with phred
v0.980904 c. Vector identified by cross_match with the database 18
and animals 12 options
PCR primers
FORWARD: AGGAACAGGATGATGAT
REVERSE: GTTTCACAGTACAGG
Plates: 74. Row: 0 column: 21
Seq primer: ATTAGGTGACACTATAG.
FEATURES
source
1..465
/organism="Sus scrofa"
/db_xref="taxon:9823"
/clone_lib="MARE LP15"
/issue_type="Pooled"
/lab_hosts="BHLB"
/note="Vector: pcwv smpt6; site_1: XbaI; site_2: XbaI;
Library made from pooled tissue from day 11, 13, 15, 20,
and 30 embryos."
BASE COUNT
105 a 120 c 126 g 114 t

Query Match 85.7%; Score 18; DB 164; Length 465;
Best Local Similarity 100.0%; Pred. No. 5.9e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ggaagagcgtctatctgca 19
|||||
DB 126 GGAAGAGCGTCTATCTGCA 143
RESULT 9
BE710007/c
LOCUS BE710007 508 bp mRNA EST 02-JAN-2001
DEFINITION MI-P-AVI-nqa-c-10-0-01-S1 MI-P-AVI Sus scrofa cDNA clone
KEYWORDS BE710007
EST. BE710007.1 GI:12009484
SOURCE Sus scrofa
ORGANISM Sus scrofa
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
AUTHORS 1 (bases 1 to 508)
TITLE Bonaldo, M.F., Lennon, G., and Soares, M.B.
Normalization and subtractions: two approaches to facilitate gene
discovery
JOURNAL Genome Res. 6 (9), 791-806, (1996)
COMMENT Contact: Tuzi, CK
9704477
Molecular Genetics Laboratory, Department of Animal Science
Iowa State University
201 Kiddie Hall, Ames, IA 50011-3150, USA
Tel: 515/2944252
Fax: 515/2942401
Email: tuzi@iastate.edu
The sequence contained an oligo-dT track that was present in the
oligonucleotide that was used to prime the synthesis of first
strand cDNA and therefore this may represent a bonafide poly A

tail. The sequence tag present in the cDNA between the NotI site
and the oligo-dT track served to verify it as a clone from the
normalized placenta library preparation. M.B. Soares
Lab, University of Iowa EST sequencing: M.B. Soares Lab, University
of Iowa Clone distribution: clones will be available through
Research Genetics (www.resgen.com)
Seq primer: M13 Forward
POLYA Yes.

FEATURES
source
1..508
/organism="Sus scrofa"
/strain="crossbreed"
/db_xref="taxon:9823"
/clone_lib="MI-P-AVI"
/clone_nqa-c-10-0-01"
/lab_hosts="BHLB (Life Technologies)"
/note="Vector: pTZ19-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: EcoRI; The MI-P-AVI
library is normalized library derived from the MI-P-AVI
library, ultimately derived from placenta tissue, for a
detailed description of the library from which this clone
was derived, please visit our web site at
http://plquest.genome.iastate.edu/. The procedure used to
create this library has been previously described (Bonaldo
, Lennon and Soares, Genome Research 6: 791-806, 1996)
TAG_LIB=MI-P-AVI
TAG_TISSUE=placenta
TAG_SEQ=ATTG3"
BASE COUNT 117 a 135 c 141 g 125 t

Query Match 85.7%; Score 18; DB 168; Length 508;
Best Local Similarity 100.0%; Pred. No. 5.9e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ggaagagcgtctatctgca 19
|||||
DB 359 GGAAGAGCGTCTATCTGCA 442
RESULT 10
BE387918
LOCUS BE387918 511 bp mRNA EST 21-JUL-2000
DEFINITION 601282166F1 NIH_MGC_44 Homo sapiens cDNA clone IMAGE:4604808 5';
KEYWORDS BE387918
EST. BE387918.1 GI:9333284
SOURCE human.
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS NIH MGC Int'l. <http://www.ncbi.nlm.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Tel: (401) 496-1550
Email: Robert_Strausberg@nih.gov
FISSE Procurement: AICP
cDNA Library Preparation: Ling Hong/Klein Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNI)
DNA Sequencing by: Invivo Genomics, Inc.
clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNI at:
<http://image.llni.gov>
Plate: LLCM249 row: d column: 17
High quality sequence stop: 459.

FEATURES
source
1..511
/organism="Homo sapiens"
/db_xref="taxon:9606"

Query Match 100.0%; Score 21; DB 1; Length 1002;
 Best Local Similarity 100.0%; Pred. No. 0.45;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 aaaaagagcgcctacacatg 21
 |||
 DB 488 AGGAGAGCGCTGGGCGATG 508

RESULT 2

US-09-345-409-1
 : Sequence 3, Application US/00445409
 : Patent No. 5121029
 : GENERAL INFORMATION:
 : APPLICANT: Schnupp, Thomas
 : APPLICANT: Lison, James
 : APPLICANT: Molnar, Istvan
 : APPLICANT: Zirkle, Ross
 : APPLICANT: Cyt, Devon
 : APPLICANT: Goelach, Joern
 : TITLE OF INVENTION: GENES FOR THE RE-SYNTHESIS OF PESTICIDES
 : FILE REFERENCE: 4-30582A
 : CURRENT APPLICATION NUMBER: US/09/345,409
 : CURRENT FILING DATE: 1999-06-17
 : NUMBER OF SEQ ID NOS: 40
 : SOFTWARE: Patent In Ver. 2.0
 : SEQ ID NO 1
 : LENGTH: 68750
 : TYPE: DNA
 : ORGANISM: Soranquium cellulosum
 US-09-345-409-1

Query Match 73.4%; Score 15.4; DB 4; Length 48750;
 Best Local Similarity 94.1%; Pred. No. 1.1e+02;
 Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 aaaaagagcgcctacacatg 17
 |||
 DB 41747 aaaaagagcgcctacacatg 31764

RESULT 3

US-09-346-844-49-1
 : Sequence 45, Application US/0044634
 : Patent No. 6156857
 : GENERAL INFORMATION:
 : APPLICANT: Ratph, David
 : APPLICANT: Aki, Gao
 : APPLICANT: Ohara, Mark S.
 : APPLICANT: Vellei, Robert
 : TITLE OF INVENTION: DIAGNOSIS OF DISEASE STATE USING mRNA
 : TITLE OF INVENTION: PROFILES IN PLETHRAL LIQUOR
 : NUMBER OF SEQUENCES: 55
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: Arnold, White & Durkee
 : STREET: P.O. Box 4433
 : CITY: Houston
 : STATE: Texas
 : COUNTRY: USA
 : ZIP: 77210
 : MEDIUM TYPE: Floppy disk
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: PC-DOS/MS-DOS
 : SOFTWARE: Patent In Release #1.0, Version #1.40
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/09/346,894
 : FILING DATE: Concurrently Herewith
 : CLASSIFICATION:
 : PRIOR APPLICATION DATA:

: APPLICATION NUMBER: US 60/041,576
 : FILING DATE: 24-MAR-1997
 : ATTORNEY/AGENT INFORMATION:
 : NAME: Nakashima, Richard A.
 : REGISTRATION NUMBER: 42,024
 : REFERENCE/BOOKLET NUMBER: 09/001,014
 : TELECOMMUNICATION INFORMATION:
 : TELEPHONE: (512) 418-3000
 : TELEFAX: (512) 474-7577
 : INFORMATION FOR SEQ ID NO: 49:
 : SEQUENCE CHARACTERISTICS:
 : LENGTH: 183 base pairs
 : TYPE: nucleic acid
 : STRANDEDNESS: single
 : TOPOLOGY: linear
 US-09-046-894-49

Query Match 72.4%; Score 15.2; DB 4; Length 183;
 Best Local Similarity 85.0%; Pred. No. 1.3e+02;
 Matches 17; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 aggaagagcgcctacacatg 20
 |||
 DB 110 AGGAGAGCGCTGGGCGATG 91

RESULT 4

US-08-485-229-3
 : Sequence 3, Application US/08485229
 : Patent No. 5605690
 : GENERAL INFORMATION:
 : APPLICANT: Jacobs, Cindy A.
 : APPLICANT: Smith, Craig A.
 : TITLE OF INVENTION: Method of Treating TNF-independent
 : TITLE OF INVENTION: Inflammation Using Tumor Necrosis Factor Antagonists
 : NUMBER OF SEQUENCES: 5
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: Immunex Corporation
 : STREET: 51 University Street
 : CITY: Seattle
 : STATE: Washington
 : COUNTRY: U.S.A.
 : ZIP: 98101
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: 5.25 floppy disk
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: PC-DOS/MS-DOS
 : SOFTWARE: Patent In Release #1.0, Version #1.25
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/08/485,229
 : FILING DATE:
 : CLASSIFICATION: 435
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: US/07/944,246
 : FILING DATE:
 : ATTORNEY/AGENT INFORMATION:
 : NAME: Wight, Christopher L.
 : REGISTRATION NUMBER: 41,680
 : REFERENCE/BOOKLET NUMBER: 256;
 : TELECOMMUNICATION INFORMATION:
 : TELEPHONE: (206) 587-0430
 : TELEFAX: (206) 587-0606
 : INFORMATION FOR SEQ ID NO: 3:
 : SEQUENCE CHARACTERISTICS:
 : LENGTH: 1557 base pairs
 : TYPE: nucleic acid
 : STRANDEDNESS: single
 : TOPOLOGY: linear
 : MOLECULE TYPE: cDNA
 : HYPOTHETICAL: NO
 : ANTI-SENSE: NO
 : IMMEDIATE SOURCE:

? TELEPHONE: (212) 977-9550
? TELEFAX: (212) 664-0525
? TELEX: (212) 42523 COOP UI
? INFORMATION FOR SEQ ID NO: 1:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 2140 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: single
? TOPOLOGY: unknown
? MOLECULE TYPE: DNA (genomic)
? HYPOTHEICAL: N
? ANTI-SENSE: N
? FEATURE:
? NAME/KEY: CDS
? LOCATION: 178..1893
? OTHER INFORMATION:
US-08-228-932-1

Query Match. 70.5% Score 14.8; DB 1; Length 2140;
Best Local Similarity 88.9%; Pred. No. 26+02;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 ggaagcgcctgagcggat 20
|||||
Db 1660 GGAAGCGCTGGTGGCT 1643

RESULT 12
US-08-468-939-1/c
? Sequence 1, Application US/08468939
? Patent No. 5714381
? GENERAL INFORMATION:
? APPLICANT: Jonathan A. Bard et al.
? TITLE OF INVENTION: DNA Encoding Human Alpha 1 Adrenergic
? TITLE OF INVENTION: Receptors and Uses Thereof
? NUMBER OF SEQUENCES: 6
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: COOPER & DUNHAM LLP
? STREET: 1185 Avenue of the Americas
? CITY: New York
? STATE: New York
? COUNTRY: U.S.A.
? ZIP: 10036
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: PatentIn Release #1.24
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/468,939
? FILING DATE:
? CLASSIFICATION: 435
? ATTORNEY/AGENT INFORMATION:
? NAME: White, John P.
? REGISTRATION NUMBER: 28,678
? REFERENCE/DOCKET NUMBER: 41337 A
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (212) 278-0400
? TELEFAX: (212) 391-0526
? INFORMATION FOR SEQ ID NO: 1:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 2140 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: single
? TOPOLOGY: unknown
? MOLECULE TYPE: DNA (genomic)
? HYPOTHEICAL: N
? ANTI-SENSE: N
? FEATURE:
? NAME/KEY: CDS
? LOCATION: 178..1893
? OTHER INFORMATION:

US-08-468-939-1

Query Match. 70.5% Score 14.8; DB 1; Length 2140;
Best Local Similarity 88.9%; Pred. No. 26+02;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 ggaagcgcctgagcggat 20
|||||
Db 1660 GGAAGCGCTGGTGGCT 1643

RESULT 13
US-08-406-855A-1/c
? Sequence 1, Application US/08406855A
? Patent No. 5861309
? GENERAL INFORMATION:
? APPLICANT: Jonathan A. Bard et al.
? TITLE OF INVENTION: DNA Encoding Human Alpha 1 Adrenergic
? TITLE OF INVENTION: Receptors and Uses Thereof
? NUMBER OF SEQUENCES: 23
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Cooper & Dunham LLP
? STREET: 1185 Avenue of the Americas
? CITY: New York
? STATE: New York
? COUNTRY: U.S.A.
? ZIP: 10036
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: PatentIn Release #1.30
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/406,855A
? FILING DATE: 21-AUG-1995
? CLASSIFICATION: 435
? ATTORNEY/AGENT INFORMATION:
? NAME: White, John P.
? REGISTRATION NUMBER: 28,678
? REFERENCE/DOCKET NUMBER: 41337 A
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (212) 278-0400
? TELEFAX: (212) 391-0526
? INFORMATION FOR SEQ ID NO: 1:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 2140 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: single
? TOPOLOGY: unknown
? MOLECULE TYPE: DNA (genomic)
? HYPOTHEICAL: N
? ANTI-SENSE: N
? FEATURE:
? NAME/KEY: CDS
? LOCATION: 178..1893
? OTHER INFORMATION:
US 08 406-855A-1

Query Match. 70.5% Score 14.8; DB 2; Length 2140;
Best Local Similarity 88.9%; Pred. No. 26+02;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 ggaagcgcctgagcggat 20
|||||
Db 1660 GGAAGCGCTGGTGGCT 1643

RESULT 14
US-08-722-190-1/c
? Sequence 1, Application US/08722190
? Patent No. 5990128


```

FI      exon      /number 5
FI      /*tag n
FI      /number 7
FI      /note- "Although the specification states that the
FI      exon/intron boundary is between position 4966 and
FI      4997 it is believed to be at position 4955 as a
FI      splice at position 4996-4997 alters the
FI      translation of the protein represented in W36509"
FI      intron     4997..5305
FI      /*tag o
FI      /number 7
FI      /number 7
FI      /*tag p
FI      /number 8
FI      /number 5498
FI      /*tag q
FI      /number 8
FI      /*tag r
FI      /number 9
FI      /number 5733
FI      /*tag s
FI      /number 9
FI      /*tag t
FI      /number 10
FI      /*tag u
FI      /number 10
FI      /*tag v
FI      /number 11
FI      /number 6319
FI      /*tag w
FI      /number 11
FI      /number 6395
FI      /*tag x
FI      /number 12
FI      /number 6510
FI      /*tag y
FI      /number 12
FI      /number 7083
FI      /*tag z
FI      /number 13
FI      /note "Although the specification states that this
FI      exon/intron boundary is between position 7698 and
FI      7699 there is a putative intron within this
FI      exon at position 7236 to 7284. the sliced
FI      out of this intron allows a reading frame which
FI      translates to the protein represented in W36509"
FI      intron     7699..8943
FI      /*tag aa
FI      /number 14
FI      /number 8944
FI      /*tag ab
FI      /number 14
FI      /number 9063
FI      /*tag ac
FI      /number 14
FI      /number 9428
FI      /*tag ad
FI      /number 15
FI      /number 9484
FI      /*tag ae
FI      /number 15
FI      /number 9956
FI      /*tag af
FI      /number 16
FI      /number 10101
FI      /*tag ag
FI      /number 16
FI      /number 10471

```

```

FI      intron     /*tag ah
FI      /number 17
FI      /note "Although the specification states that this
FI      exon/intron boundary is between position
FI      10644 and 10645 it is believed to be at
FI      position 10650 to 10651 as a splice at
FI      position 10644 to 10645 results in the loss of
FI      the Val and Gly residues at positions 849 and
FI      850 of the protein W36509"
FI      intron     10645..10742
FI      /*tag ai
FI      /number 17
FI      /number 10743
FI      /*tag aj
FI      /number 18
FI      /number 10827
FI      /*tag ak
FI      /number 18
FI      /number 11287
FI      /*tag al
FI      /number 19
FI      /number 11449
FI      /*tag am
FI      /number 19
FI      /number 11287
FI      /*tag an
FI      /number 19
FI      /number 11449
FI      /*tag ao
FI      /number 19
FI      /number 11529
FI      /*tag ap
FI      /number 20
FI      /number 11745
FI      /*tag aq
FI      /number 20
FI      /number 11974
FI      /*tag ar
FI      /number 21
XX      W09740855-A1.
XX      06 NOV-1997.
XX      01 OCT-1996; 96W0-0815769.
XX      29 APR 1996; 96JC 3016482.
XX      (OF) J UNIV JOHN HOPKINS SCHOOL MEDICINE.
XX      Dietz BV.
XX      W011 1997-549494/50.
XX      P-PSDB; W36509.
XX      New isolated regulators of non sense mediated RNA decay - used to
XX      develop products for the study, diagnosis and therapy of disorders
XX      such as Marfan Syndrome, accelerated ageing and cancers
XX      Claim 1; Fig 4; 79pp; English.
XX      this sequence encodes the murine RNU1 protein which regulates
XX      telomerase mediated RNA decay (2000). The RNU1 (regulator of non-sense
XX      transcripts) protein and other products can be used in the study,
XX      diagnosis and therapy of disorders involving NMRs such as Marfan
XX      Syndrome, accelerated ageing or various cancers.
XX      Sequence 13146 BP; 2863 A; 3390 G; 671 G; 3222 T; 0 other;

```

```

Query Match: 47.9%; Score 24; DB 18; Length 13146;
Best Local Similarity 74.4%; Prod. No. 17;
Matches 29; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

```



```
Query Match      100.0%; Score 48; DB 1; Length 1002;
Best Local Similarity 100.0%; Pred. No. 4,7e-19;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 aaactaaagatctcaaacaccccaaacacaaagatctcaaacacaa 48
      |||||
DB 554 AAAATCATAGATCTCAAAACACCAACCAACCAAGGTTCACAAACAGACA 601

RESULT 2
US-08 791-495 1
: Sequence 1, Application US/08791495
: Patent No. 5811519
: GENERAL INFORMATION:
: APPLICANT: Leth, Bernard
: APPLICANT: Lucas, Sophie
: APPLICANT: De Smet, Charles
: APPLICANT: O'Boyle, Pauline
: APPLICANT: Boon-Fallier, Thierry
: TITLE OF INVENTION: LI-1 TUMOR SPECIFIC GENES
: NUMBER OF SEQUENCES: 14
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Wolf, Greenfield & Sackel, P.C.
: STREET: 600 Atlantic Avenue
: CITY: Boston
: STATE: MA
: COUNTRY: USA
: ZIP: 02210
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent In Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08,791,495
: FILING DATE:
: CLASSIFICATION: 4,35
: ATTORNEY/AGENT INFORMATION:
: NAME: Van Amerdam, John R.
: REGISTRATION NUMBER: 40,212
: REFERENCE TO OTHER SEQUENCES: US/08,791,495
: TELEPHONE: 617-720-2441
: TELEFAX: 617-720-2441
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 217 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: HYPOTHETICAL: NO
: ANTI-SENSE: NO
: FRAGMENT TYPE: internal
: ORIGINAL SOURCE:
: ORGANISM: Homo sapiens
US-08 791-495-1

Query Match      81.2%; Score 39; DB 1; Length 217;
Best Local Similarity 100.0%; Pred. No. 8.1e-07;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 gattcacaacacacacacaaagatctcaaacacaa 48
      |||||
DB 1 GATCTCAGACATCTCAAAACACCAACCAAGGTTCACAAACAGACA 49

RESULT 3
US-08 724-354D-3/c
: Sequence 3, Application US/08724354/c
: Patent No. 6048965
: GENERAL INFORMATION:
: APPLICANT: Dietz, Harry C.
: TITLE OF INVENTION: MAMMALIAN REGULATOR OF
: NUMBER OF SEQUENCES: 29
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Fish & Richardson, P.C.
: STREET: 4225 Executive Square, Suite 1400
: CITY: La Jolla
: STATE: CA
: COUNTRY: USA
: ZIP: 92037
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: Windows95
: SOFTWARE: FastSeq for Windows Version 2.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08,724,354D
: FILING DATE: 01-OCT-1996
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 60/016,482
: FILING DATE: 29-APR-1996
: ATTORNEY/AGENT INFORMATION:
: NAME: Hailo, Lisa A.
: REGISTRATION NUMBER: 48,347
: REFERENCE TO OTHER SEQUENCES: US/08,724,354D-1
: TELEPHONE: 619-678-5070
: TELEFAX: 619-678-5099
: INFORMATION FOR SEQ ID NO: 3:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 13146 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: genomic DNA
US-08-724-354D-3

Query Match      47.9%; Score 23; DB 2; Length 13146;
Best Local Similarity 74.4%; Pred. No. 3.4;
Matches 29; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 6 taaagatctcaaacacacacaaagatctcaaacaa 44
      |||
DB 4429 TAGGAATCTATACAGCCTAACCTAATTAATCAACACCA 3491

RESULT 4
US-09-270-984A-3/c
: Sequence 3, Application US/09270984A
: Patent No. 6048965
: GENERAL INFORMATION:
: APPLICANT: Dietz, Harry C.
: TITLE OF INVENTION: MAMMALIAN REGULATOR OF
: NUMBER OF SEQUENCES: 29
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Fish & Richardson, P.C.
: STREET: 4225 Executive Square, Suite 1400
: CITY: La Jolla
: STATE: CA
: COUNTRY: USA
: ZIP: 92037
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: Windows95
: SOFTWARE: FastSeq for Windows Version 2.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US,09,270,984A
```



```

1  FILLING DATE: 27 SEP 1991
2  PRIOR APPLICATION DATA:
3  APPLICATION NUMBER: US 07/080,441
4  FILLING DATE: 7 SEP 1990
5  PRIOR APPLICATION DATA:
6  APPLICATION NUMBER: US 07/080,442
7  FILLING DATE: 20 JUN 1989
8  PRIOR APPLICATION DATA:
9  APPLICATION NUMBER: US 07/029,018
10 FILLING DATE: 24 MAR 1989
11 PRIOR APPLICATION DATA:
12 APPLICATION NUMBER: US 08/045,957
13 FILLING DATE: 12 APR 1994
14 ATTORNEY/AGENT INFORMATION:
15 NAME: Elmer, James Scott
16 REGISTERED NUMBER: 46,129
17 REFERENCE: 1,124 base pairs
18 TELEPHONE: (919) 541-8614
19 TELEFAX: (919) 541-8609
20 INFORMATION FOR SEQ. ID NO. 46:
21 SEQUENCE CHARACTERISTICS:
22 LENGTH: 1,124 base pairs
23 TYPE: nucleic acid
24 STRANDEDNESS: single
25 TOPOLOGY: linear
26 MOLECULE TYPE: DNA (genomic)
27 ORGANISM: Cucumber satyris
28 INDIVIDUAL INFORMATION:
29 NAME: Cucumber satyris
30 STRAIN: Cucumber satyris
31 SEX: male
32 AGE: 1 day
33 DATE OF BIRTH: 1991
34 DATE OF DEATH: 1991
35 DATE OF COLLECTION: 1991
36 DATE OF ANALYSIS: 1991
37 DATE OF DEPOSIT: 1991
38 DATE OF RELEASE: 1991
39 DATE OF REVISION: 1991
40 DATE OF REVISION: 1991
41 DATE OF REVISION: 1991
42 DATE OF REVISION: 1991
43 DATE OF REVISION: 1991
44 DATE OF REVISION: 1991
45 DATE OF REVISION: 1991
46 DATE OF REVISION: 1991
47 DATE OF REVISION: 1991
48 DATE OF REVISION: 1991
49 DATE OF REVISION: 1991
50 DATE OF REVISION: 1991
51 DATE OF REVISION: 1991
52 DATE OF REVISION: 1991
53 DATE OF REVISION: 1991
54 DATE OF REVISION: 1991
55 DATE OF REVISION: 1991
56 DATE OF REVISION: 1991
57 DATE OF REVISION: 1991
58 DATE OF REVISION: 1991
59 DATE OF REVISION: 1991
60 DATE OF REVISION: 1991
61 DATE OF REVISION: 1991
62 DATE OF REVISION: 1991
63 DATE OF REVISION: 1991
64 DATE OF REVISION: 1991
65 DATE OF REVISION: 1991
66 DATE OF REVISION: 1991
67 DATE OF REVISION: 1991
68 DATE OF REVISION: 1991
69 DATE OF REVISION: 1991
70 DATE OF REVISION: 1991
71 DATE OF REVISION: 1991
72 DATE OF REVISION: 1991
73 DATE OF REVISION: 1991
74 DATE OF REVISION: 1991
75 DATE OF REVISION: 1991
76 DATE OF REVISION: 1991
77 DATE OF REVISION: 1991
78 DATE OF REVISION: 1991
79 DATE OF REVISION: 1991
80 DATE OF REVISION: 1991
81 DATE OF REVISION: 1991
82 DATE OF REVISION: 1991
83 DATE OF REVISION: 1991
84 DATE OF REVISION: 1991
85 DATE OF REVISION: 1991
86 DATE OF REVISION: 1991
87 DATE OF REVISION: 1991
88 DATE OF REVISION: 1991
89 DATE OF REVISION: 1991
90 DATE OF REVISION: 1991
91 DATE OF REVISION: 1991
92 DATE OF REVISION: 1991
93 DATE OF REVISION: 1991
94 DATE OF REVISION: 1991
95 DATE OF REVISION: 1991
96 DATE OF REVISION: 1991
97 DATE OF REVISION: 1991
98 DATE OF REVISION: 1991
99 DATE OF REVISION: 1991
100 DATE OF REVISION: 1991

```


GenCore version 4.5
Copyright (c) 1996-2000 CompuLink Ltd.

EM function: nucleic search, using SW model

Run on: May 1, 2001, 08:03:58 : Search time: 3786.75 Seconds
(without alignments)
110.757 Million cell updates/sec

Title: US-09-341-829A-4_Copy_554_601

Perfect search:

Sequence:

Search label: HENILLY_NUP

Gap: 10.0, Gapext: 1.0

Scattered: 962517 seqs, 40604920 residues

Total number of hits satisfying chosen parameters: 19247034

Minimum Lib seq length: 1

Maximum Lib seq length: 20000000

Post processor: Minimum Match: 99

Maximum Match: 100

Listed: 11131 45 summaries

Database:

EST:

1: db_est1:
2: db_est2:
3: db_est3:
4: db_est4:
5: db_est5:
6: db_est6:
7: db_est7:
8: db_est8:
9: db_est9:
10: db_est10:
11: db_est11:
12: db_est12:
13: db_est13:
14: db_est14:
15: db_est15:
16: db_est16:
17: db_est17:
18: db_est18:
19: db_est19:
20: db_est20:
21: db_est21:
22: db_est22:
23: db_est23:
24: db_est24:
25: db_est25:
26: db_est26:
27: db_est27:
28: db_est28:
29: db_est29:
30: db_est30:
31: db_est31:
32: db_est32:
33: db_est33:
34: db_est34:
35: db_est35:
36: db_est36:
37: db_est37:
38: db_est38:
39: db_est39:
40: db_est40:
41: db_est41:
42: db_est42:
43: db_est43:
44: db_est44:
45: db_est45:
46: db_est46:
47: db_est47:

44: em_esthum10:
45: em_esthum11:
46: em_esthum12:
47: em_esthum13:
48: em_esthum14:
49: em_esthum15:
50: em_esthum16:
51: em_esthum17:
52: em_esthum18:
53: em_esthum19:
54: em_esthum20:
55: em_esthum21:
56: em_esthum22:
57: em_esthum23:
58: em_esthum24:
59: em_esthum25:
60: em_esthum26:
61: em_esthum27:
62: em_esthum28:
63: em_esthum29:
64: em_esthum30:
65: em_esthum31:
66: em_esthum32:
67: em_esthum33:
68: em_esthum34:
69: em_esthum35:
70: em_esthum36:
71: em_esthum37:
72: em_esthum38:
73: em_esthum39:
74: em_esthum40:
75: em_esthum41:
76: em_esthum42:
77: em_esthum43:
78: em_esthum44:
79: em_esthum45:
80: em_esthum46:
81: em_esthum47:
82: em_esthum48:
83: em_esthum49:
84: em_esthum50:
85: em_esthum51:
86: em_esthum52:
87: em_esthum53:
88: em_esthum54:
89: em_esthum55:
90: em_esthum56:
91: em_esthum57:
92: em_esthum58:
93: em_esthum59:
94: em_esthum60:
95: em_esthum61:
96: em_esthum62:
97: em_esthum63:
98: em_esthum64:
99: em_esthum65:
100: em_esthum66:
101: em_esthum67:
102: db_est25:
103: db_est26:
104: db_est27:
105: db_est28:
106: db_est29:
107: db_est30:
108: db_est31:
109: db_est32:
110: db_est33:
111: db_est34:
112: db_est35:
113: db_est36:
114: db_est37:
115: db_est38:
116: db_est39:

117: qb_est48: *
118: qb_est49: *
119: qb_est50: *
120: qb_est51: *
121: qb_est52: *
122: qb_est53: *
123: qb_est54: *
124: qb_est55: *
125: qb_est56: *
126: qb_est57: *
127: qb_est58: *
128: qb_est59: *
129: qb_est60: *
130: qb_est61: *
131: qb_est62: *
132: qb_est63: *
133: qb_est64: *
134: qb_est65: *
135: qb_est66: *
136: qb_est67: *
137: qb_est68: *
138: qb_est69: *
139: qb_est70: *
140: qb_est71: *
141: qb_est72: *
142: qb_est73: *
143: qb_est74: *
144: qb_est75: *
145: qb_est76: *
146: qb_est77: *
147: qb_est78: *
148: qb_est79: *
149: qb_est80: *
150: qb_est81: *
151: qb_est82: *
152: qb_est83: *
153: qb_est84: *
154: qb_est85: *
155: qb_est86: *
156: qb_est87: *
157: qb_est88: *
158: qb_est89: *
159: qb_est90: *
160: qb_est91: *
161: qb_est92: *
162: qb_est93: *
163: qb_est94: *
164: qb_est95: *
165: qb_est96: *
166: qb_est97: *
167: qb_est98: *
168: qb_est99: *
169: qb_est100: *
170: qb_est101: *
171: qb_est102: *
172: qb_est103: *
173: qb_est104: *
174: qb_est105: *
175: qb_est106: *
176: qb_est107: *
177: qb_est108: *
178: qb_est109: *
179: qb_est110: *
180: qb_est111: *
181: qb_est112: *
182: qb_est113: *
183: qb_est114: *
184: qb_est115: *
185: qb_est116: *
186: qb_est117: *
187: qb_est118: *
188: qb_est119: *
189: qb_est120: *

190: em_qss_pln1: *
191: em_qss_pln2: *
192: em_qss_pro: *
193: em_qss_rnd1: *
194: em_qss_rnd2: *
195: em_qss_rnd3: *
196: em_qss_rnd4: *
197: em_qss_rnd5: *
198: em_qss_vrt1: *
199: em_qss_vrt2: *
200: em_qss_vrt3: *
201: qb_qss1: *
202: qb_qss2: *
203: qb_qss3: *
204: qb_qss4: *
205: qb_qss5: *
206: qb_qss6: *
207: qb_qss7: *
208: qb_qss8: *
209: qb_qss9: *
210: qb_qss10: *
211: qb_qss11: *
212: qb_qss12: *
213: qb_qss13: *
214: qb_qss14: *
215: qb_qss15: *
216: qb_qss16: *
217: qb_qss17: *
218: qb_qss18: *
219: qb_qss19: *
220: qb_qss20: *
221: qb_qss21: *
222: qb_qss22: *
223: qb_qss23: *
224: qb_qss24: *
225: qb_qss25: *
226: qb_qss26: *
227: qb_qss27: *
228: qb_qss28: *
229: qb_qss29: *
230: qb_qss30: *
231: qb_qss31: *
232: qb_qss32: *
233: qb_qss33: *
234: qb_qss34: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result	Query			DB	ID	Description
	No.	Score	Match			
1	48	100.0	559	165	BE276194	BE276194 601144442
2	48	100.0	586	166	BE390259	BE390259 601285131
3	48	100.0	689	165	BE274123	BE274123 601120646
4	35.4	73.8	739	166	BE391232	BE391232 601285820
5	30.6	63.8	800	166	BE408892	BE408892 601303769
6	28.2	58.8	674	166	BE388562	BE388562 601281855
c 7	26	54.2	400	151	BE662245	BE662245 maa00003
c 8	26	54.2	415	148	BE452889	BE452889 maa00003
c 9	26	54.2	475	146	BE321797	BE321797 626410937
c 10	26	54.2	534	137	BE635044	BE635044 uv760073
c 11	26	54.2	542	137	BE635044	BE635044 uv760073
c 12	26	54.2	579	3	AA163654	AA163654 int19q11.r
c 13	26	54.2	686	173	BE081278	BE081278 H0050008
c 14	25.8	52.8	526	225	AZ637601	AZ637601 1M0497M03
c 15	25.8	53.8	640	218	AZ284687	AZ284687 R001 23-1
c 16	25.2	53.8	135	238	AA157332	AA157332 R001 23-1
c 17	25.2	52.5	467	193	AA180099	AA180099 01318043
c 18	25.2	52.5	688	220	AZ081853	AZ081853 1M0116A11


```

DEFINITION Homo sapiens chromosome 16 clone RP11-291F1, LOW-PASS SEQUENCE
SAMPLING.
ACCESSION AC010547
VERSION V010547.2 GI:6758795
KEYWORDS HIG: HIGS_PHASE0
SOURCE human.
ORGANISM Homo Sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 214911)
AUTHORS DeJ. Joint Genome Institute.
TITLE Sequencing of Human Chromosome 16
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 214911)
AUTHORS DeJ. Joint Genome Institute.
TITLE Direct Submission
JOURNAL Submitted (15-SEP-1999) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
on Jan 26, 2000 this sequence version replaced 41:5382401.
COMMENT -----Genome Center
Center: Joint Genome Institute
Center Code: JGI
Web Site: http://www.jgi.doe.gov
-----
* NOTE: this record contains 422 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.
1 365: contig of 365 bp in length
366 gap of unknown length
1137 contig of 771 bp in length
1138 gap of unknown length
1139 contig of 794 bp in length
1140 gap of unknown length
1941 contig of 676 bp in length
2607 gap of unknown length
3285 contig of 679 bp in length
3286 gap of unknown length
3874 contig of 588 bp in length
3875 gap of unknown length
4171 contig of 298 bp in length
4172 gap of unknown length
4173 contig of 741 bp in length
4174 gap of unknown length
4913 contig of 727 bp in length
4914 gap of unknown length
5640 contig of 258 bp in length
5641 gap of unknown length
5642 contig of 562 bp in length
5643 gap of unknown length
6460 contig of 969 bp in length
6461 gap of unknown length
7429 contig of 598 bp in length
7430 gap of unknown length
8027 contig of 664 bp in length
8028 gap of unknown length
8691 contig of 749 bp in length
8692 gap of unknown length
9440 contig of 613 bp in length
9441 gap of unknown length
10053 contig of 442 bp in length
10495 contig of 887 bp in length
11482 contig of 249 bp in length
11483 gap of unknown length
12456: contig of 826 bp in length
12457 gap of unknown length
12775: contig of 319 bp in length
12776 gap of unknown length
13306: contig of 531 bp in length
13307 gap of unknown length
13407 contig of 182 bp in length
13408 gap of unknown length
13489 contig of 161 bp in length
13490 gap of unknown length
13650 contig of 778 bp in length
13651 gap of unknown length
15252: contig of 125 bp in length
15253 gap of unknown length
15986: contig of 234 bp in length
15987 gap of unknown length
16216: contig of 730 bp in length
16217 gap of unknown length
17500: contig of 784 bp in length
17501 gap of unknown length
18374: contig of 874 bp in length
18375 gap of unknown length
19470: contig of 996 bp in length
19471 gap of unknown length
19458: contig of 488 bp in length
19459 gap of unknown length
20049: contig of 191 bp in length
20050 gap of unknown length
20844: contig of 795 bp in length
20845 gap of unknown length
21641: contig of 797 bp in length
21642 gap of unknown length
22125: contig of 484 bp in length
22126 gap of unknown length
22482: contig of 177 bp in length
22483 gap of unknown length
23005: contig of 703 bp in length
23006 gap of unknown length
23709: contig of 704 bp in length
23710 gap of unknown length
24895: contig of 1186 bp in length
24896 gap of unknown length
25012: contig of 117 bp in length
25013 gap of unknown length
25853: contig of 841 bp in length
25854 gap of unknown length
26726: contig of 873 bp in length
26727 gap of unknown length
27416: contig of 690 bp in length
27417 gap of unknown length
28134: contig of 718 bp in length
28135 gap of unknown length
28255: contig of 121 bp in length
28256 gap of unknown length
28513: contig of 258 bp in length
28514 gap of unknown length
29227: contig of 714 bp in length
29228 gap of unknown length
29692: contig of 465 bp in length
29693 gap of unknown length
30843: contig of 1151 bp in length
30844 gap of unknown length
31551: contig of 708 bp in length
31552 gap of unknown length
32314: contig of 763 bp in length
32315 gap of unknown length
32437: contig of 123 bp in length
32438 gap of unknown length
32795: contig of 358 bp in length
32796 gap of unknown length
32972: contig of 177 bp in length
32973 gap of unknown length
33261: contig of 289 bp in length
33262 gap of unknown length

```


TITLE
JOURNAL
 Submitted (03-MAY-2000) Genome Sequencing Center, Washington
 University School of Medicine, 444 Forest Park Parkway, St. Louis,
 Mo 63108, USA
REFERENCE
AUTHORS
TITLE
JOURNAL
 4 (bases 1 to 37566)
 Watson,R.H.
 Direct Submission
 Submitted (10-AUG-2000) Genome Sequencing Center, Washington
 University School of Medicine, 444 Forest Park Parkway, St. Louis,
 Mo 63108, USA
 5. (bases 1 to 37566)
 Watson,R.H.
 Direct Submission
 Submitted (07-SEP-2000) Genome Sequencing Center, Washington
 University School of Medicine, 444 Forest Park Parkway, St. Louis,
 Mo 63108, USA
 6. (bases 1 to 37566)
 Watson,R.H.
 Direct Submission
 Submitted (07-OCT-2000) Department of Genetics, Washington
 University, 444 Forest Park Avenue, St. Louis, Missouri 63108, USA
 7. Aug 30, 2000 this sequence version replaced q1:2606665.
 ----- Genome Center
 Center: Washington University Genome Sequencing Center
 Center code: WUGSC
 Web site: http://pilot.wustl.edu/
 Contact: sapient@watson.wustl.edu
 ----- Summary Statistics

 Center project name: HNH024309

NOTICE: This sequence may not represent the entire insert of this
 clone. It may be shorter because we only sequenced overlapping
 clone sections once, or longer because we provide a small overlap
 between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
 all regions were double stranded, sequenced with an alternate
 chemistry or covered by high quality PCR (100% phage quality
 40); an attempt was made to resolve all sequencing problems, such
 as compressions and repeats; all regions were checked by a second
 from more than one subclone; and the assembly was confirmed by
 restriction digest.

MAPEING INFORMATION:
 The position of this clone was established as part of a
 collaboration between the Human Chromosome Y Mapping Project
 (Tomoko Kawauchi, Helen Skaletsky, Laura G. Brown, Steve Rozen,
 and David C. Page at the Whitehead Institute for Biomedical
 Research, Cambridge MA) and the Washington University Genome
 Sequencing Center, St. Louis MO.

SOURCE INFORMATION:
 The RPCL11 human BAC library was made from the blood of one male
 donor, as described by Gossawa, K., Wren, P., Zhang, P., Freese, E.,
 Iatano, M., Catanesi, J.J., and de Jong, P.J. (1998) An improved
 approach for construction of bacterial artificial chromosome
 libraries. Genomics 51:1-8. The clone may be obtained either from
 Research Genetics, Inc. (http://www.regen.com) or from the
 and coworkers at the Roswell Park Cancer Institute
 (http://cancer.roswell.org/).

FEATURES: PRACE3,6

RECHIBING SOURCE INFORMATION:
 The clone sequenced to the left is RP11-945E12; the clone sequenced
 to the right is RP11-56A33. Actual start of this clone is at base
 position 1 of RP11-2433P.

FEATURES
 source
 Location/Qualifiers
 1..37566
 /base14444444 "Name sapiens"
 /db_xref "taxon:9606"
 /chromosome:"Y"
 /map:"Y"
 /clone "RP11:2433P"

/clone_lib "RPCL11"
 20..100
 /rpt_family "CRI"
 102..161
 /rpt_family "ERV1"
 707..775
 /rpt_family "L2"
 1014..1486
 /rpt_family "MgLR"
 1720..2008
 /rpt_family "Alu"
 2013..2551
 /rpt_family "L1"
 2655..3516
 /rpt_family "L1"
 3537..3705
 /rpt_family "ERV1"
 4448..4775
 /rpt_family "ERV"
 4777..4884
 /rpt_family "ERV"
 5930..5245
 /rpt_family "Alu"
 6265..6342
 /rpt_family "MER1_type"
 7503..7494
 /rpt_family "MER2_type"
 8385..8608
 /rpt_family "Alu"
 9437..9388
 /rpt_family "L2"
 10109..10247
 /rpt_family "Alu"
 10365..10498
 /rpt_family "MER"
 10642..10789
 /rpt_family "similarity" to EST AF197971 (NID:06862555) "
 10662..10962
 /rpt_family "similarity" to EST AA812229 (NID:042879588) od14b05.sl"
 10704..10789
 /rpt_family "similarity" to EST 221236 (NID:027975) "
 10938..11101
 /rpt_family "01"
 11163..11350
 /rpt_family "similarity" to EST AF197971 (NID:06862555) "
 11168..11345
 /rpt_family "similarity" to EST 221236 (NID:027975) "
 11189..11345
 /rpt_family "similarity" to EST AA812229 (NID:042879588) od14b05.sl"
 11726..11827
 /rpt_family "similarity" to EST AF197971 (NID:06862555) "
 12484..12794
 /rpt_family "Alu"
 14198..14283
 /rpt_family "ERV1"
 14812..15247
 /rpt_family "L1"
 16074..16160
 /rpt_family "L2"
 16515..16810
 /rpt_family "Alu"
 19142..19277
 /rpt_family "Alu"
 19278..19441
 /rpt_family "Alu"
 19603..19966
 /rpt_family "L1"
 19970..20057
 /rpt_family "Alu"
 20028..20421
 /rpt_family "L1"
 20600..20736
 /rpt_family "ERV1"

RESULT 6
 ID Z49049 standard: DNA; 1773 BP.
 XX
 AC Z49049.
 XX
 DT 31-MAR-2000 (first entry)
 XX
 DE HCV RNA polymerase coding sequence.
 XX
 KW RNA polymerase, HCV, infection, therapy; ss.
 XX
 OS Hepatitis C virus.
 XX
 PN W09967396-A1.
 XX
 PD 29-SEP-1996
 XX
 PF 24-JUN-1995; 99WO-JP03381.
 XX
 PR 24-JUN-1998; 98JP 0177817.
 XX
 PA (ITRE-) INT REAGENTS CORP.
 PA (CHOS-) CHUGAI SEIYAKU KK
 PA (TOYO-) TOYODA T.
 XX
 PI Toyoda T, Kohara M, Kohara K, Higashi K, Tsuchiya M.
 XX
 DR WPI: 2000-106297/00
 DR p-bpna, y6a576
 XX
 PI Polynucleotides encoding a hepatitis C virus (HCV)-derived RNA
 PI polymerase, useful for screening polymerase inhibitors to treat HCV -
 PS Disclosure; Page 14 17, 25pp, Japanese.
 XX
 CC This sequence encodes the hepatitis C virus (HCV) RNA polymerase protein
 CC of the invention. The protein can be used to detect inhibitors of RNA
 CC polymerase, for use in the treatment of hepatitis C viral infections
 XX
 SQ Sequence 1773 BP; 412 A; 515 C; 461 G; 385 T, 0 other

Query Match 43.9%; Score 25; DB 21; Length 1773;
 Best Local Similarity 64.9%; Pred. No. 3.8;
 Matches 37; Conservative 0; Mismatches 20; Indels 0; Gaps 0.
 QY 1 ccacccgaaggaagccaggaagatgggtgcagaagatgaccccttaagtaatttc 57
 DB 1483 ccaccttgagactgagacatcggcccaagatgtccgcgtacgtctgccc 1539

RESULT 7
 Q33282
 ID Q33282 standard: cDNA; 9472 BP.
 XX
 AC Q33282;
 XX
 DT 19-MAY-1993 (first entry)
 XX
 DE Korean hepatitis C virus full cDNA sequence LNC1.
 XX
 KW HCV-LNC1; diagnosis; vaccine; ds.
 XX
 OS Korean hepatitis C virus.
 XX
 FH Key Location/Qualifiers
 FT CDS 343..9375
 FT /*tag< a
 FT /feature= "HCV polyprotein"
 XX

PN EP521318-A.
 XX
 PD 07-JAN-1993.
 XX
 DT 10 JUN 1992, 92EP 0109753.
 XX
 PR 10-JUN-1991; 91KP-0009510.
 PR 06-AUG-1991; 91KP-0013601.
 XX
 PA (LUCK-) LUCKY LTD.
 XX
 PI Cho JM, Choi DY, Kim CH, Kim ST, Lee YB, Lim KJ, Park YW;
 PI So HS, Yang JY;
 XX
 DR WPI: 1993-001843/01
 DR p-PSDB; R30616.
 XX
 FT DNA and polypeptide(s) from a new type of hepatitis C virus (HCV)
 FT for diagnosing and vaccinating against HCV infections
 XX
 PS Claim 7; Fig 2; 119pp; English.
 XX
 CC This is the full cDNA sequence of Korean hepatitis C virus (HCV)
 CC cDNA, HCV LNC1. Portions of the cDNA may be used as probes or
 CC primers to diagnose the presence of the virus in putative samples.
 CC Polypeptides derived from the cDNA may be used in a specific and
 CC accurate method for detecting HCV antibodies in the serum of
 CC hepatitis C patients. Antibodies directed against these polypeptides
 CC are useful for the purification of HCV antigens and for the
 CC development of an improved diagnostic to detect HCV antigens in a
 CC sample. The polypeptides may also be used in a vaccine for treatment
 CC and prevention of HCV infection at a dosage of 5-200 ug/peptide.
 XX
 SQ Sequence 9472 BP; 1897 A; 2829 C; 2396 G; 2073 T; 0 other;

Query Match 43.9%; Score 25; DB 14; Length 9472;
 Best Local Similarity 64.9%; Pred. No. 5.3;
 Matches 37; Conservative 0; Mismatches 23; Indels 0; Gaps 0.
 QY 1 ccacccgaaggaagccaggaagatgggtgcagaagatgaccccttaagtaatttc 57
 DB 5829 ccaccttgagactgagacatcggcccaagatgtccgcgtacgtctgccc 9139

RESULT 8
 P12467/c
 ID P12467 standard: cDNA; 703 BP.
 XX
 AC P12467;
 XX
 DT 13-MAR-2001 (first entry)
 XX
 DE Aspergillus oryzae EST SEQ ID NO.4990.
 XX
 KW Multiple gene expression; filamentous fungal cell; EST;
 KW expressed sequence tag, Fusarium venenatum, Aspergillus niger;
 KW Aspergillus oryzae; trichoderma reesei; identification; recombination;
 KW culture condition, environmental stress; spore morphogenesis;
 KW metabolic pathway engineering; catabolic pathway engineering; ss.
 XX
 OS Aspergillus oryzae.
 XX
 PN W0200056762-A2.
 XX
 PD 28-SEP-2000
 XX
 PF 22-MAR-2000; 2000WO-US07781.
 XX
 PR 22-MAR-1999; 99US-0273623.
 XX
 PA (NOVO) NOVO NORDISK BIOTECH INC.
 PA (NOVO) NOVO NORDISK AS.


```

PF 17-SEP-1990; 90EP-0310149.
XX
PR 21-DEC-1989; 89US-0456142.
PR 15-SEP-1989; 89US-0408045.
XX
PA (CHIK) CHIRON CORP.
PA (OYAA/) OYA A.
XX
PI Cha T., Ban J., Houghton M., Irvine BD., Kolberg JA;
PI Miyamura T., Saito I., Weiner AJ;
XX
DR WPI: 1999-48843/41.
DR P-PSDB; Y14974.
XX
PI New Hepatitis C Virus isolates, useful for diagnosis of hepatitis
PI infections and development of vaccines
XX
PS Example 2; Fig 5; 122pp; English.
XX
CC The invention provides two new isolates of hepatitis C virus (HCV), J1
CC and J7. These two isolates comprise nucleotide and amino acid sequences
CC that are distinct from the HCV isolate HCV-1. The nucleotide sequences
CC may be used to detect non-A, non-B HCV (NANBH) polynucleotides by
CC hybridisation for diagnosis of NANBH infections. They may also be used to
CC screen blood donors, donated blood and blood products for this infection.
CC The isolates may also be used to isolate other naturally occurring
CC variants of the virus. The polypeptides may be used as a vaccine for
CC administration to patients to protect against infection with NANBH. The
CC present sequence represents the J1 NS5 domain coding sequence.
XX
SQ Sequence 427 BP; 91 A; 129 C; 105 G; 99 T; 3 other;

```

```

Query Match 41.1%; Score 22.4; DB 20; Length 427;
Best Local Similarity 63.2%; Prod. No. 11;
Matches 36; Conservative 0; Mismatches 21; Indels 0; Gaps 0;
QY 1 ccaaccgaaggagcccaagagatggatgcagagatgcttcaatqatgttc 57
DB 345 ccaaccttcaagatctggagacatcgggccagaagatgcgcgctaaagctactatcc 391

```

Search completed: May 1, 2001, 04:00:18
 Job time: 7123 sec

1 LENGTH: 4933 base pairs
2 TYPE: nucleic acid
3 STRANDEDNESS: single
4 TOP-LOGY: linear
5 MOLECULE TYPE: cDNA to mRNA
6 HYPOTHETICAL: NO
7 ANTI-SENSE: NO
8 ORIGINAL SOURCE:
9 ORGANISM: Homo sapiens
10 CELL TYPE: mucosal lymphocyte
11 FEATURE:
12 NAME/KEY: CDS
13 LOCATION: 126..4662
14 FEATURE:
15 NAME/KEY: mat_peptide
16 LOCATION: 180..3659
17 FEATURE:
18 NAME/KEY: sig_peptide
19 LOCATION: 126..179
20 US-08-199-776-1

Query Match 49.38; Score 22.4; DB 1; Length 3933;
Best Local Similarity 66.7%; Pred. No. 11;
Matches 32; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 5 ccaagaaagcccaagaaatgaatgaagatgaacacatttaataaa 52
DB 2128 cccagagtcggcagagcagatgtcgggaagggcggggcactaataatca 2081

RESULT 8
US-08-663-731-1/c
1 Sequence 1, Application US/08664731
2 Patent No. 6057423
3 GENERAL INFORMATION:
4 APPLICANT: Brenner, Michael P.
5 APPLICANT: Parker, Christina M.
6 TITLE OF INVENTION: No. 6057423c1 integrin alpha subunit
7 NUMBER OF SEQUENCES: 25
8 CORRESPONDENCE ADDRESS:
9 ADDRESSEE: Well, Greenfield and Sacks, P.C.
10 STREET: 600 Atlantic Avenue
11 CITY: Boston
12 STATE: MA
13 COUNTRY: USA
14 ZIP: 02210
15 COMPUTER READABLE FORM:
16 MEDIUM TYPE: Floppy disk
17 COMPUTER: IBM PC compatible
18 OPERATING SYSTEM: PC-DOS/MS-DOS
19 SOFTWARE: Patent in Release #1.0, Version #1.25
20 CURRENT APPLICATION DATA: 08/199,776
21 APPLICATION NUMBER: 08/199,776
22 FILING DATE:
23 CLASSIFICATION:
24 PRIOR APPLICATION DATA:
25 FILING DATE:
26 ATTORNEY/AGENT INFORMATION:
27 NAME: Fliemer, Elizabeth R.
28 REGISTRATION NUMBER: 36,637
29 REFERENCE WORKS: MIMPPF: 6091/7020
30 TELECOMMUNICATION INFORMATION:
31 TELEPHONE: 617-720-9592
32 TELEFAX: 617-720-2441
33 INFORMATION FOR SEQ ID NO: 1:
34 SEQUENCE CHARACTERISTICS:
35 LENGTH: 4933 base pairs
36 TYPE: nucleic acid
37 STRANDEDNESS: single
38 TOP-LOGY: linear
39 MOLECULE TYPE: cDNA to mRNA

1 HYPOTHETICAL: NO
2 ANTI-SENSE: NO
3 ORIGINAL SOURCE:
4 ORGANISM: Homo sapiens
5 CELL TYPE: mucosal lymphocyte
6 FEATURE:
7 NAME/KEY: CDS
8 LOCATION: 126..4662
9 FEATURE:
10 NAME/KEY: mat_peptide
11 LOCATION: 180..3659
12 FEATURE:
13 NAME/KEY: sig_peptide
14 LOCATION: 126..179
15 US-08-663-731-1

Query Match 49.38; Score 22.4; DB 3; Length 3933;
Best Local Similarity 66.7%; Pred. No. 11;
Matches 32; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 5 ccaagaaagcccaagaaatgaatgaagatgaacacatttaataaa 52
DB 2128 cccagagtcggcagagcagatgtcgggaagggcggggcactaataatca 2081

RESULT 9
US-08-879-338-1/c
1 Sequence 1, Application US/08879338A
2 Patent No. 6064906
3 GENERAL INFORMATION:
4 APPLICANT: Brenner, Michael P.
5 APPLICANT: Parker, Christina M.
6 TITLE OF INVENTION: Antibodies to No. 6064906e1 integrin Alpha
7 TITLE OF INVENTION: Submit
8 FILE REFERENCE: 6091/7080/ERP
9 CURRENT APPLICATION NUMBER: 08/36,879,338A
10 CURRENT FILING DATE: 1997-06-20
11 EARLIER APPLICATION NUMBER: 08/36,879,338A
12 EARLIER FILING DATE: 1996-06-14
13 EARLIER APPLICATION NUMBER: 08/199,776
14 EARLIER FILING DATE: 1994-02-18
15 NUMBER OF SEQ ID NOS: 31
16 SOFTWARE: FastSeq for Windows Version 3.0
17 SEQ ID NO 1
18 LENGTH: 3933
19 TYPE: DNA
20 ORGANISM: Homo Sapiens
21 FEATURE:
22 NAME/KEY: CDS
23 LOCATION: (126) ... (4659)
24 NAME/KEY: mat_peptide
25 LOCATION: (180) ... (3659)
26 NAME/KEY: sig_peptide
27 LOCATION: (126) ... (179)
28 US-08-879-338-1

Query Match 49.38; Score 22.4; DB 3; Length 3933;
Best Local Similarity 66.7%; Pred. No. 11;
Matches 32; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 5 ccaagaaagcccaagaaatgaatgaagatgaacacatttaataaa 52
DB 2128 cccagagtcggcagagcagatgtcgggaagggcggggcactaataatca 2081

RESULT 10
PCT-US95-02044-1/c
1 Sequence 1, Application PCT/US9502044
2 GENERAL INFORMATION:
3 APPLICANT:
4 TITLE OF INVENTION: Novel integrin alpha subunit


```
1 APPLICATION NUMBER: US 2000/023,452
2 FILING DATE: 08-SEP-1995
3 ATTORNEY/AGENT INFORMATION:
4 NAME: Strella, John R.
5 REGISTRATION NUMBER: 42,944
6 REFERENCE/WORK NUMBER: 915499-00122005
7 TELECOMMUNICATION INFORMATION:
8 TELEPHONE: (415) 576-0200
9 TELEFAX: (415) 576-0300
10 INFORMATION FOR SEQ ID NO: 1:
11 SEQUENCE CHARACTERISTICS:
12 LENGTH: 4232 base pairs
13 TYPE: nucleic acid
14 STRANDEDNESS: single
15 ORIENTATION: linear
16 MOLECULE TYPE: cDNA
17 FEATURE:
18 NAME/KEY: CDS
19 LOCATION: 1..4315
20 OTHER INFORMATION: /product: "TPC2"
21 FEATURE:
22 NAME/KEY: unsure
23 LOCATION: 651
24 OTHER INFORMATION: /note "N" probably C"
25 FEATURE:
26 NAME/KEY: unsure
27 LOCATION: 660
28 OTHER INFORMATION: /note "N" - probably C"
29 FEATURE:
30 NAME/KEY: unsure
31 LOCATION: 1142
32 OTHER INFORMATION: /note "N" - probably T"
33 FEATURE:
34 NAME/KEY: unsure
35 LOCATION: 1150
36 OTHER INFORMATION: /note "N" - maybe G"
37 FEATURE:
38 NAME/KEY: unsure
39 LOCATION: 1188
40 OTHER INFORMATION: /note "N" - probably G"
41 FEATURE:
42 NAME/KEY: unsure
43 LOCATION: 1211
44 OTHER INFORMATION: /note "N" - probably C"
45 FEATURE:
46 NAME/KEY: unsure
47 LOCATION: 1225
48 OTHER INFORMATION: /note "N" - probably G"
49 FEATURE:
50 NAME/KEY: unsure
51 LOCATION: 1230
52 OTHER INFORMATION: /note "N" probably G"
53 FEATURE:
54 NAME/KEY: unsure
55 LOCATION: 1248
56 OTHER INFORMATION: /note "N" - probably C"
57 FEATURE:
58 NAME/KEY: unsure
59 LOCATION: 1247
60 OTHER INFORMATION: /note "N" probably A"
61 FEATURE:
62 NAME/KEY: unsure
63 LOCATION: 1301
64 OTHER INFORMATION: /note "N" - probably G"
65 FEATURE:
66 NAME/KEY: unsure
67 LOCATION: 1375
68 OTHER INFORMATION: /note "N" - probably C"
69 FEATURE:
70 NAME/KEY: unsure
71 LOCATION: 1379
72 OTHER INFORMATION: /note "N" - probably C"
73 FEATURE:
74 NAME/KEY: unsure
75 LOCATION: 1491
76 OTHER INFORMATION: /note "N" - probably C"
77 FEATURE:
78 NAME/KEY: unsure
79 LOCATION: 1407
80 OTHER INFORMATION: /note "N" - probably C"
81 FEATURE:
82 NAME/KEY: unsure
83 LOCATION: 1540
84 OTHER INFORMATION: /note "N" - probably G"
85 FEATURE:
86 NAME/KEY: unsure
87 LOCATION: 1543
88 OTHER INFORMATION: /note "N" - probably G"
89 FEATURE:
90 NAME/KEY: unsure
91 LOCATION: 1545
92 OTHER INFORMATION: /note "N" - probably G"
93 FEATURE:
94 NAME/KEY: unsure
95 LOCATION: 1586
96 OTHER INFORMATION: /note "N" - probably G"
97 FEATURE:
98 NAME/KEY: unsure
99 LOCATION: 1588
100 OTHER INFORMATION: /note "N" - probably G"
101 FEATURE:
102 NAME/KEY: unsure
103 LOCATION: 1652
104 OTHER INFORMATION: /note "N" - probably G"
105 FEATURE:
106 NAME/KEY: unsure
107 LOCATION: 1685
108 OTHER INFORMATION: /note "N" - probably C"
109 FEATURE:
110 NAME/KEY: unsure
111 LOCATION: 1688
112 OTHER INFORMATION: /note "N" - probably G"
113 FEATURE:
114 NAME/KEY: unsure
115 LOCATION: 1707
116 OTHER INFORMATION: /note "N" - probably C"
117 FEATURE:
118 NAME/KEY: unsure
119 LOCATION: 1719
120 OTHER INFORMATION: /note "N" - probably C"
121 FEATURE:
122 NAME/KEY: unsure
123 LOCATION: 1725
124 OTHER INFORMATION: /note "N" - probably C"
125 FEATURE:
126 NAME/KEY: unsure
127 LOCATION: 1724
128 OTHER INFORMATION: /note "N" - probably C"
129 FEATURE:
130 NAME/KEY: unsure
131 LOCATION: 1789
132 OTHER INFORMATION: /note "N" - probably G"
133 FEATURE:
134 NAME/KEY: unsure
135 LOCATION: 1796
136 OTHER INFORMATION: /note "N" - probably A"
137 FEATURE:
138 NAME/KEY: unsure
139 LOCATION: 1816
140 OTHER INFORMATION: /note "N" - probably G"
141 FEATURE:
142 NAME/KEY: unsure
143 LOCATION: 1924
144 OTHER INFORMATION: /note "N" - probably G"
145 FEATURE:
146 NAME/KEY: unsure
```



```

117: qb_est48:*
118: qb_est49:*
119: qb_est50:*
120: qb_est51:*
121: qb_est52:*
122: qb_est53:*
123: qb_est54:*
124: qb_est55:*
125: qb_est56:*
126: qb_est57:*
127: qb_est58:*
128: qb_est59:*
129: qb_est60:*
130: qb_est61:*
131: qb_est62:*
132: qb_est63:*
133: qb_est64:*
134: qb_est65:*
135: qb_est66:*
136: qb_est67:*
137: qb_est68:*
138: qb_est69:*
139: qb_est70:*
140: qb_est71:*
141: qb_est72:*
142: qb_est73:*
143: qb_est74:*
144: qb_est75:*
145: qb_est76:*
146: qb_est77:*
147: qb_est78:*
148: qb_est79:*
149: qb_est80:*
150: qb_est81:*
151: qb_est82:*
152: qb_est83:*
153: qb_est84:*
154: qb_est85:*
155: qb_est86:*
156: qb_est87:*
157: qb_est88:*
158: qb_est89:*
159: qb_est90:*
160: qb_est91:*
161: qb_est92:*
162: qb_est93:*
163: qb_est94:*
164: qb_est95:*
165: qb_est96:*
166: qb_est97:*
167: qb_est98:*
168: qb_est99:*
169: qb_est100:*
170: qb_est101:*
171: qb_est102:*
172: qb_est103:*
173: qb_est104:*
174: qb_est105:*
175: qb_est106:*
176: qb_est107:*
177: qb_est108:*
178: qb_est109:*
179: qb_est110:*
180: qb_est111:*
181: qb_est112:*
182: qb_est113:*
183: qb_est114:*
184: qb_est115:*
185: qb_est116:*
186: qb_est117:*
187: qb_est118:*
188: qb_est119:*
189: qb_est120:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result	No.	Score	Query Match	Length	DB	ID	Description
	1	47.6	83.5	589	165	BE274124	BE274124 601120646
	2	37.4	65.6	800	166	BE408892	BE408892 601304769
	3	32.8	57.5	729	166	BE391232	BE391232 601385920
	4	27.7	47.1	1141	171	BE242629	BE242629 602377118
c	5	26.2	46.0	923	15	AT068383	AT068383 meta0002a
c	6	25.6	44.9	440	14	A1094027	A1094027 601161018
c	7	25.6	44.9	588	221	A2411525	A2411525 1M0:84007
c	8	25.6	44.9	641	1	AA047981	AA047981 mj26e10.r
c	9	25.6	44.9	964	240	CNS020B4	AL214248 Tetraodon
c	10	25.6	44.9	1061	232	CNS051A9	AL342321 Tetraodon
c	11	25.2	44.2	831	134	AW350275	AW350275 60210007B
c	12	24.9	43.5	1073	232	CNS05279	AL119542 Tetraodon
c	13	24.5	43.2	798	171	BE366127	BE366127 602286341
c	14	24.4	42.8	118	118	AW670670	AW670670 114815 MA
c	15	24.4	42.8	354	227	H48579	H48579 RPO111-2118
c	16	24.2	42.5	123	223	A2558246	A2558246 RPO111-2118
c	17	24.2	42.5	590	148	BE703444	BE703444 60238N111
c	18	24.2	42.5	955	150	BE575748	BE575748 602145474

Query Match:	46.00%	Score: 26.2	100	15	Length 924
Best local Similarity:	69.40%	Prod. No.	56		
Matches:	44	Conserved Sites	0	Mismatches	15
				Indels	0
				Gaps	0

cellular, at least contains NF- κ B³⁶ in ilomoliprin cDNA clone (MW: 62,047 kD) similar to contains Alu repetitive element containing element MER2 repetitive element³⁷, mRNA sequence.

NATIONAL
 CANCER
 INSTITUTE
 NATIONAL CANCER INSTITUTE, Cancer Genome Anatomy Project (CGAP),
 Library Index
 Published (1997)
 Cancer Research Strategic
 Planning

1996b). Consent form (Info image, Ind. spec) for further information.
 Image length: 5x1 Std Error: 0.00
 Gen primer: primer name ambiguous

a modified polylinker; Site 1: Not 1; Site 2: Eco RI; equal amounts of plasmid DNA from three normalized libraries (total lung NBHL9W, testis NHF, and B cell

was PCR amplified cDNAs from pools of 5,000 viruses made from the same 5 libraries. The pools consisted of 1 M.A.V. F, 100,000, 297,400 + 30,2107, 692,632 + 697,249,

AZ411525.1 G110545.09
 VERSION
 KEYWORDS
 GSS,
 HOUSE MOUSE,
 SOURCE

1 (Issues 1 to 566)

REFERENCE

AUTHORS

DAVID, D., AGYAL, A., BARBER, M., BOWEN, I., LOVELL, R., HAMILL, C.,
ESTAM, H., LAMBERS, S., MATHIAS, M., MORGAN, E., PETERSEN, R., RAY, J.

Journal of
Plasma Medicine
Copyright © 2000
Lippincott Williams & Wilkins
Contact: Robert B. Weiss

84112, USA
Tel: 801 580 6406
Fax: 801 580 7177

EXEMPLES
 Seq primer: 5'-TACACACAAAGAGAGATATGAGC-3'
 class: plasmid ends
 High quality sequence stop: 0000.
 Location: 200,000,000 bp

Journal of the American Chemical Society
Vol. 117, No. 1, pp. 1-10, 1995
Copyright 1995 by the American Chemical Society
0887-624X/95/0001-0001\$05.00/0
DOI: 10.1021/ja00117a001

polynucleotide kinase. Adapter oligonucleotides were ligated to the blunt ends in high molar excess. The adapter RNA was purified and size selected on a 4% to

inducible derivative of pMS2-E1. The vector was limited with adaptors complementary to the insert adaptors and purified. The sheared, adaptor-mouse tRNA was immediately

BASE COUNT	160 d	141 c	133 g	154 f
ORIGIN				

scale clone and sequencing project of the tetraodon nigroviridis genome. For more information, please take a look at <http://www.genoscope.cng.fr/tnetraodon>

```

RASP COUNT      269 3    242 0    312 4    235 4    16 08475
/cont_000.D
/note %genoscope sequence ID : C0AB020H05C1 end : T7"

```

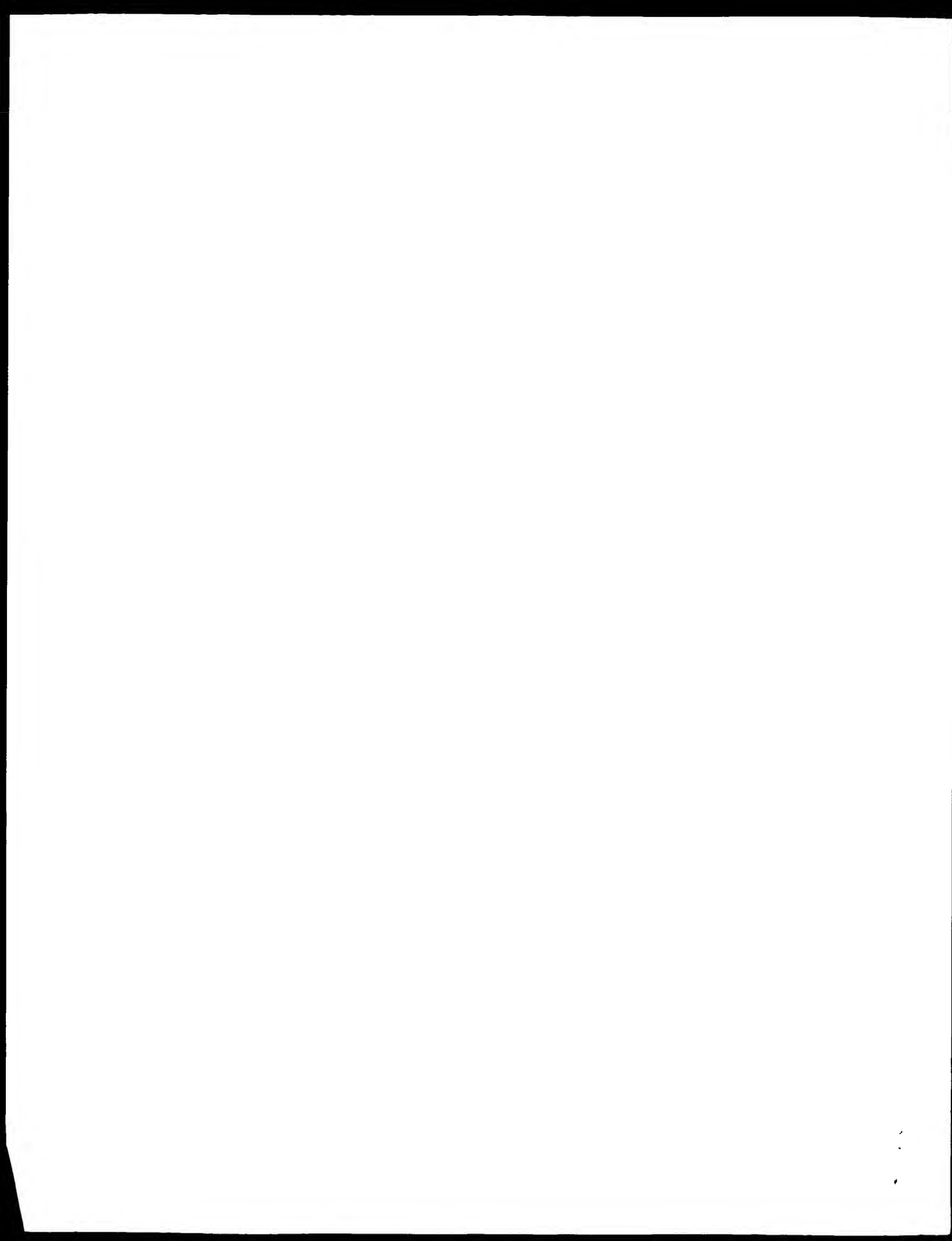
ORGANISM	SUS SCROFA
CLASSIFICATION	Eukaryota; Metazoa; Chordata; Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE	1 (bases 1 to 148)
AUTHORS	Fahrenkrug, S.C., Fieking, B.A., Bohrer, G.A., Smith, T.P., Lucas, F., Stone, R.T., Houston, M.P., Grosso, W.M., Bennett, G.A., Loughrey, W.W., and Keele, J.W.

Email: smith@mail.marco.usda.gov
 Fax: 402.762.4390
 Single-pass sequencing. Bases called and aligned with paired
 v0.560904.e. Vectors identified by crossmatch with the miniscope
 and mismatch 12 options.
 18K Primers:
 FORWARD: AGGAACAGCTATGACAT
 BACKWARD: GTTTCGAGTCAGCAGG
 Plate: 44 row: N column: 5
 Seq primer: ATTATGACACATATAG

```

Query Match          42.8%; Score 24.4; DB 118; Length 148;
Best Local Similarity 69.0%; Pred. No. 1.8e+02;
Matches 34; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

```

GENERAL INFORMATION:
APPLICANT: Ieth , Bernard
APPLICANT: Lucas, Sophie
APPLICANT: De Smet, Charles
APPLICANT: Godelaine, Daniele
APPLICANT: Boon-Failler, Thierry
TITLE OF INVENTION: IL-1 TUMOR SPECIFIC GENES
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Wolf, Greenfield & Sacks, P.C.
STREET: 600 Atlantic Avenue
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/06/791,495
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Van Amsterdam, John R.
REGISTRATION NUMBER: 40,212
REFERENCE/BOOKLET NUMBER: L0461/0005
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-720-3500
TELEFAX: 617-720-2441
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 19 nucleotides
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: YES
ORIGINAL SOURCE:
ORGANISM: Homo sapiens

50 70 80 90 100 110 120
 GTCTCTCCAGACACAAATCAGGCTGCTACAAAGAACCTGGCCATCTGCTGCTAGGG
 GATCTCAGACACA-CCCAATCAG
 x 10 20

140
GAGCAGCATGACCTAG GAGCGGTATGATTTTCTTAACTTTCTCTTAAGAT CTTCAGCC
||| || | | | | | | | | | | | | | | | |
AAGTCTCACACACAGACCTTGATCACTAGATGCTGGCCATGATGAGAGGCC--AGGAGATGGGTGCA
30 40 50 60 70 80 90

260
 AAAAAACCGGCGAGAAAGCACTG---CGTGTATCCATTCACACAT-CTAAAG-CTGCTCCACACACACCTGGA
 |||
 210
 GAAGTGTGTCCTTTAAATGATGTCTCTGCGCCCTCATATTCGCCCATCTGCTGCGAGACACCCAGCCCAA
 |||
 100
 110 120 130 140 150 160 170 180 190 200 210 220 230 240 250 260

270
 - TGGAGAGCTGTCAGTGTGGGTGGTGTCTTATAGCATCTGACCTGGGTTAAATCTGACGAGGACAGACATCACA...T
 310
 320
 330
 CTGGAGGCTCTTCA-TCAGCTGCTGTCTGACACAG-CTTTC-CTGTTGATGGGATC
 170 180 190 200 210 X
 340 350 360 370
 AAAAGCGACAGCTCTGACGACCATCTCCCTGGGGCTC

5. US-04-341-829A-4' (1-1002)
US-08-791-495-11 Sequence 11, Application US/08791495
Sequence 11, Application US/08791495

Initial Score	=	19	Optimized Score	=	19	Significance	=	-0.45
Residue Identity	=	100%	Matches	=	19	Mismatches	=	0
Gaps	=	0	Conservative Substitutions	=	0			0

CATGTTACCGCTGGTCTTGTCAGCAGCAAAACAATCAGAGCCGCATAAAGAATGGCCACTGTGTGTGGCACATA
CTGACCACATGTGTCTGGGA

120 130 140 150 160
TTCCTAGCGGAGGAGGATGACCTAGGAATGGGTGCTCAGGCTGGCC

6. US 09 341-829A-4' (1-1002)
US-08-791-495-3 Sequence 3, Application US/08791495

Sequence 3, Application US98791495
GENERAL INFORMATION:
APPLICANT: Leth, Bernard
APPLICANT: Lucas, Sophie
APPLICANT: De Smet, Charles
APPLICANT: Godelaine, Daniele
APPLICANT: Boon-Fallier, Thierry
TITLE OF INVENTION: LL-1 TUMOR SPECIFIC GENES
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESS: Wm. Greenfield & Sons, P.C.
STREET: 600 Atlantic Avenue
CITY: Boston



FASTA File Information

FASTA File pairwise Comparison of Sequences

Results File us-09-341-829a-4.res made by sdsaid on Tue 1 May 101 9:24:50 PDT.

Query sequence being compared: US 09-341-829A 4 (1-1002)

Number of sequences searched: 11

Number of sequences above cutoff: 11

Results of the initial comparison of US 09-341-829A 4 (1-1002) with:

File: US09341495.seq

	100	90	80	70	60	50	40	30	20	10	0
N	1	1	1	1	1	1	1	1	1	1	1
Y	1	1	1	1	1	1	1	1	1	1	1
M	1	1	1	1	1	1	1	1	1	1	1
P	1	1	1	1	1	1	1	1	1	1	1
R	1	1	1	1	1	1	1	1	1	1	1
T	1	1	1	1	1	1	1	1	1	1	1
C	1	1	1	1	1	1	1	1	1	1	1
G	1	1	1	1	1	1	1	1	1	1	1
A	1	1	1	1	1	1	1	1	1	1	1
S	1	1	1	1	1	1	1	1	1	1	1
SCORE	111	223	334	445	557	668	779	891	1002		
SDRV	0	0	0	0	0	0	0	0	0	0	0

PARAMETERS

Similarity matrix	Unitary	K Updo	4
Mismatch penalty	1	Joining penalty	40
Gap penalty	1.00	Window size	32
Gap size penalty	0.33		
Cutoff score	1		
Randomization group	9		

SEARCH STATISTICS

Search	Mean	Median	Standard deviation
	298	19	426.45
Time:	cpu		Total elapsed
	00:00:00.00		00:00:00.00
Number of residues:			2958
Number of sequences searched:			11
Number of sequences above cutoff:			11

The scores below are sorted by initial score.

Similarity is calculated based on initial score.

A 100% identical sequence to the query sequence was found:

Sequence Name	Description	Length	Score	Sim. Frame
---------------	-------------	--------	-------	------------

1. US-08-791-495-4 Sequence 4, Application US 1002 1002 1002 2.45 0

The list of other best scores is:

Sequence Name	Description	Length	Score	Sim. Frame
2. US-08-791-495-6 Sequence 6, Application US		255	518	574 0
3. US-08-791-495-8 Sequence 8, Application US		795	474	540 0
4. US-08-791-495-1 Sequence 1, Application US		217	217	217 0
5. US-08-791-495-12 Sequence 12, Application		18	18	18 0
6. US-08-791-495-10 Sequence 10, Application		18	18	18 0
7. US-08-791-495-2 Sequence 2, Application US		18	18	18 0
8. US-08-791-495-14 Sequence 14, Application		20	9	11 0
9. US-08-791-495-13 Sequence 13, Application		17	8	9 0
10. US-08-791-495-3 Sequence 3, Application US		19	8	13 0
11. US-08-791-495-11 Sequence 11, Application		19	7	11 0

1. US-09-341-829A 4 (1-1002)
US-08-791-495-4 Sequence 4, Application US/08791495

Sequence 4, Application US/08791495

GENERAL INFORMATION:

APPLICANT: Beth. Bernard

APPLICANT: Lucas, Sophie

APPLICANT: de Smet, Charles

APPLICANT: Godelaine, Daniele

APPLICANT: Boon-Pallour, Thierry

TITLE OF INVENTION: LL-1 1006R SPECIFIC GENES

NUMBER OF SEQUENCES: 14

CORRESPONDENCE ADDRESS:

ADDRESSEE: Wolf, Greenfield & Sacks, P.C.

STREET: 600 Atlantic Avenue

CITY: Boston

STATE: MA

COUNTRY: USA

ZIP: 02210

COMPUTER READABLE FORM:

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC DOS/MS DOS

SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: 09/08791495

FILING DATE:

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Van Amsterdam, John R.

REGISTRATION NUMBER: 40,212

REFERENCE/DECKET NUMBER: 10451/7005

TELECOMMUNICATION INFORMATION:

TELEPHONE: 617-720-3500

TELEFAX: 617-720-2441

INFORMATION FOR SEQ. ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 1002 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: cDNA

HYPOTHETICAL: NO

ANTI-SENSE: NO

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: CDS

LOCATION: 65..697

Initial Score	1002	Identical Score	1002	Similarity	2.45
Residue Identity	100%	Matches	1002	Mismatches	0
Gaps	0	Conservative Substitutions			0

[illegible]

Sequence 1, Application US/08791495

GENERAL INFORMATION:
APPLICANT: Ioth, Bernard
APPLICANT: Lucas, Sophie
APPLICANT: De Smet, Charles
APPLICANT: Godolaine, Lucile
APPLICANT: Room-fallier, Thierry
TITLE OF INVENTION: CLIPPER SPECIFIC GENES
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Wolf, Greenfield & Sacks, P.C.
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US08/791,495

FILING DATE:
 CLASSIFICATION:
 ATTORNEY/AGENT INFORMATION:
 NAME: Van Amerindam, John R.
 REGISTRATION NUMBER: 40,212
 REFERENCE/DOCKET NUMBER: J04b1/7005
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 617 720-3500
 TELEFAX: 617-720-2441
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 217 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 HYPOTHEetical: NO
 ART-GERE: NO
 FRAGMENT TYPE: Internal
 ORIGINAL SOURCE:
 ORGANISM: Homo sapiens

Initial Score	-	217	optimized Score	=	217	Significance
Residue Identity	=	100%	Matches	-	217	Mismatches
Gaps	=	0	Conservative Substitutions	=	0	
520	540	540	550	560	X	546
1GGTGGCTGGGGCTGGAGATGCGCTGGCGGAGAGGAGAGAGCTATGACATCTCAGAAATATGCAAAATAA						
550	560	560	570	580	590	600
AGGTCCTCAGAACAGAGACTGTGTACACCAAGGCTGGCGGCACCCGAGGATGATCCAGGACATGATGTGTACAG						
111	111	111	111	111	111	111
AGGCTCTCAGAAACAGAGACTGGTACACAGAGAGCGAGGCCAGGACCCAGGATAGATCCAGGACATGATGTGTACAG						
30	40	50	60	70	80	90
550	600	610	620	630	640	650
AGGTCCTCAGAACAGAGACTGTGTACACCAAGGCTGGCGGCACCCGAGGATGATCCAGGACATGATGTGTACAG						
111	111	111	111	111	111	111
AGGCTCTCAGAAACAGAGACTGGTACACAGAGAGCGAGGCCAGGACCCAGGATAGATCCAGGACATGATGTGTACAG						
30	40	50	60	70	80	90
660	670	680	690	700	710	720
GTGGTGGCTTAAATGTGATGTCTCTGGCGCTATATTTAGCGGATGATCTGTGCAGATCATATGGCAATGTG						
100	110	120	130	140	150	160
GTGGTGGCTTAAATGTGATGTCTCTGGCGCTATATTTAGCGGATGATCTGTGCAGATCATATGGCAATGTG						
170	180	190	200	210	X	
CAGCTCTCCATCAGCTCCCTGTATCCAGGAGCTTTCCTCTTGATGTGATC						
170	180	190	200	210	X	
740	740	750	760	770	780	800
CAGCTCTCCATCAGCTCCCTGTATCCAGGAGCTTTCCTCTTGATGTGATC						
170	180	190	200	210	X	
CAGCTCTCCATCAGCTCCCTGTATCCAGGAGCTTTCCTCTTGATGTGATC						
170	180	190	200	210	X	
TTTTTTTCTCAGGCTGGCTCAGGCGCAGAG						
810	820					
TTTTTTTCTCAGGCTGGCTCAGGCGCAGAG						

5. US-09 341-829A-4 (1-1002)
US-08-791-495-12 Sequence 12, Application US/08791495

Sequence 12, Application US/08791495

GENERAL INFORMATION:
 APPLICANT: Leth , Bernard
 APPLICANT: Lucas, Sophie
 APPLICANT: De Smet, Charles
 APPLICANT: Godelaine, Daniele
 APPLICANT: Boon-Fallou, Thierry
 TITLE OF INVENTION: IL-1 TUMOR SPECIFIC GENES
 NUMBER OF SEQUENCES: 14
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Wolf, Greenfield & Sacks, P.C.
 STREET: 600 Atlantic Avenue
 CITY: Boston
 STATE: MA
 COUNTRY: USA
 ZIP: 02210



Configure version 4.5
Copyright (c) 1994-2000 CompuLink Ltd.

OM protein: protein search, using SW model

Run on: May 1, 2001, 11:23:59; Search time 40.59 seconds
(without alignments)
9,858 Million cell updates/sec

Target sequence: us-09-341-829a-5_copy_142_148
Sequence: 1 RESIDUES

Scoring table: BLASTNM2

Gapop: 10, 0; Gapext: 0, 0

Searched: 590729 seqs, 5716426 residues

Total number of hits satisfying chosen parameters: 177502

Minimum hit seq length: 6

Maximum hit seq length: 40

Post processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: A (unseq_0401)*

Rank	Seq ID	Score	Length	Residues	Match
1	us-09-341-829a-5_copy_142_148	177502	40	5716426	100%
2	us-09-341-829a-5_copy_142_148	177502	40	5716426	100%
3	us-09-341-829a-5_copy_142_148	177502	40	5716426	100%
4	us-09-341-829a-5_copy_142_148	177502	40	5716426	100%
5	us-09-341-829a-5_copy_142_148	177502	40	5716426	100%
6	us-09-341-829a-5_copy_142_148	177502	40	5716426	100%
7	us-09-341-829a-5_copy_142_148	177502	40	5716426	100%
8	us-09-341-829a-5_copy_142_148	177502	40	5716426	100%
9	us-09-341-829a-5_copy_142_148	177502	40	5716426	100%
10	us-09-341-829a-5_copy_142_148	177502	40	5716426	100%
11	us-09-341-829a-5_copy_142_148	177502	40	5716426	100%
12	us-09-341-829a-5_copy_142_148	177502	40	5716426	100%
13	us-09-341-829a-5_copy_142_148	177502	40	5716426	100%
14	us-09-341-829a-5_copy_142_148	177502	40	5716426	100%
15	us-09-341-829a-5_copy_142_148	177502	40	5716426	100%
16	us-09-341-829a-5_copy_142_148	177502	40	5716426	100%
17	us-09-341-829a-5_copy_142_148	177502	40	5716426	100%
18	us-09-341-829a-5_copy_142_148	177502	40	5716426	100%
19	us-09-341-829a-5_copy_142_148	177502	40	5716426	100%
20	us-09-341-829a-5_copy_142_148	177502	40	5716426	100%
21	us-09-341-829a-5_copy_142_148	177502	40	5716426	100%
22	us-09-341-829a-5_copy_142_148	177502	40	5716426	100%

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Length	ID	Description
1	us-09-341-829a-5_copy_142_148	177502	40	5716426	Integrin subunit b
2	us-09-341-829a-5_copy_142_148	177502	40	5716426	Integrin subunit b
3	us-09-341-829a-5_copy_142_148	177502	40	5716426	Integrin subunit b
4	us-09-341-829a-5_copy_142_148	177502	40	5716426	Human cytoplasmic
5	us-09-341-829a-5_copy_142_148	177502	40	5716426	Integrin subunit b
6	us-09-341-829a-5_copy_142_148	177502	40	5716426	Human anti-thrombin
7	us-09-341-829a-5_copy_142_148	177502	40	5716426	Human anti-thrombin
8	us-09-341-829a-5_copy_142_148	177502	40	5716426	Human anti-thrombin
9	us-09-341-829a-5_copy_142_148	177502	40	5716426	Human anti-thrombin
10	us-09-341-829a-5_copy_142_148	177502	40	5716426	Fragment of human
11	us-09-341-829a-5_copy_142_148	177502	40	5716426	House dust mite al

12	24	66.7	10	14	K32448	Human thrombospondin
13	24	66.7	10	21	K02658	Human thrombospondin
14	24	66.7	19	16	K79466	Thrombospondin type
15	24	66.7	19	17	K97641	Cell binding domain
16	24	66.7	20	21	K58525	HIV-1 gp120
17	24	66.7	9	21	K44470	Human and thrombin
18	24	66.7	9	21	K44472	Human and thrombin
19	24	66.7	10	14	K25224	Residues 167-179 of
20	24	66.7	10	16	K74984	N-terminal peptide
21	24	66.7	14	22	K74446	Human class I HLA
22	24	66.7	19	17	K49202	Human leucocyte an
23	24	66.7	20	16	K80291	Trehalase release in
24	24	66.7	21	16	K80292	Trehalase release in
25	24	66.7	21	19	K20964	Human microtubule
26	24	66.7	27	19	K20971	Human neuronal lumen
27	24	66.7	30	19	K60660	Sequence specific
28	24	66.7	6	16	K71608	Sodium tuberosin
29	24	66.7	10	21	K28222	ATP binding cassette
30	24	66.7	14	17	K88440	Peptide with coil
31	24	66.7	14	17	K88445	Peptide with coil
32	24	66.7	14	17	K88446	Peptide with coil
33	24	66.7	14	20	K86171	Peptide with coil
34	24	66.7	15	16	K66421	HIV-1 TIR peptide
35	24	66.7	15	17	K88446	Peptide with coil
36	24	66.7	16	15	K88554	HIV-1 gp120 peptide
37	24	66.7	17	20	K85641	Antigenic epitope
38	24	66.7	18	11	K65990	Cell attachment pr
39	24	66.7	18	17	K88446	Peptide with coil
40	24	66.7	18	17	K82574	Alpha(v)beta(3) in
41	24	66.7	19	11	K05991	Cell attachment pr
42	24	66.7	19	19	K81022	Medwort pollen all
43	24	66.7	19	21	K65622	Peptide fragment
44	24	66.7	20	21	K88446	Alpha D peptide de
45	24	66.7	22	17	K85522	Monoclonal antibody

ANALYSIS

RESULT 1

1. First standard: peptide: 13 AA.

2. First standard: peptide: 13 AA.

3. First standard: peptide: 13 AA.

4. First standard: peptide: 13 AA.

5. First standard: peptide: 13 AA.

6. First standard: peptide: 13 AA.

7. First standard: peptide: 13 AA.

8. First standard: peptide: 13 AA.

9. First standard: peptide: 13 AA.

10. First standard: peptide: 13 AA.

11. First standard: peptide: 13 AA.

12. First standard: peptide: 13 AA.

13. First standard: peptide: 13 AA.

14. First standard: peptide: 13 AA.

15. First standard: peptide: 13 AA.

16. First standard: peptide: 13 AA.

17. First standard: peptide: 13 AA.

18. First standard: peptide: 13 AA.

19. First standard: peptide: 13 AA.

20. First standard: peptide: 13 AA.

21. First standard: peptide: 13 AA.

22. First standard: peptide: 13 AA.

23. First standard: peptide: 13 AA.

24. First standard: peptide: 13 AA.

25. First standard: peptide: 13 AA.

26. First standard: peptide: 13 AA.

27. First standard: peptide: 13 AA.

28. First standard: peptide: 13 AA.

29. First standard: peptide: 13 AA.

30. First standard: peptide: 13 AA.

31. First standard: peptide: 13 AA.

32. First standard: peptide: 13 AA.

33. First standard: peptide: 13 AA.

34. First standard: peptide: 13 AA.

35. First standard: peptide: 13 AA.

36. First standard: peptide: 13 AA.

37. First standard: peptide: 13 AA.

38. First standard: peptide: 13 AA.

39. First standard: peptide: 13 AA.

40. First standard: peptide: 13 AA.

41. First standard: peptide: 13 AA.

42. First standard: peptide: 13 AA.

43. First standard: peptide: 13 AA.

44. First standard: peptide: 13 AA.

45. First standard: peptide: 13 AA.

```

PS      Claim 3; Column 8; 8pp; English.
XX
CC      A novel cytoplasmic domain sequence (R93926) was identified in an
CC      integrin subunit beta-3 variant, termed beta-3' (see also R93929).
CC      This cytoplasmic domain or its subsequences (R93927-28) are used
CC      to produce anti-integrin antibodies useful in the analysis of
CC      cellular integrin content, in the determination of tumour origin,
CC      and in quantifying the complement of adhesion receptors on cells.
XX
SQ      Sequence 13 AA;
XX
Query Match: 77.8%; Score 28; DB 17; Length 14;
Best Local Similarity 83.3%; Pred. No. 25;
Matches 5; Conservative 1; Mismatches 0; Indels 6; Gaps 0;
XX
QY      1 REGAGR 6
DE      | : | | |
DE      2 rdaqr 7
XX
RESULT 2
R93927
ID      R93927 standard; Peptide; 17 AA.
XX
AC      R93927;
XX
DI      20-MAY-1996 (first entry)
XX
DE      Integrin subunit beta-3' cytoplasmic domain (aa13-29).
XX
KW      Integrin subunit beta-3' cytoplasmic domain; antibody; tumour.
XX
OS      Homo sapiens.
XX
PN      J5543664-A.
XX
PD      12-MAR-1996.
XX
PF      25-MAY-1989; 89DS-0357024.
XX
PR      25-MAY-1989; 89DS-0357024.
PR      09-NOV-1992; 92DS-0973547.
PR      10-MAY-1994; 94DS-0240967.
XX
PA      (LJBL) LA JOLLA CANCER RES FOUND.
XX
PI      Roostlahti EJ;
XX
DE      WPI; 1996-159740/16.
XX
PT      Peptide comprising cytoplasmic domain of integrin subunit beta-3'
PT      used to produce anti-integrin antibodies, e.g. for cell integrin
PT      content analysis for the determination tumour origin
XX
PS      Claim 1; Column 7-8; 8pp; English.
XX
CC      A novel cytoplasmic domain sequence (R93926) was identified in an
CC      integrin subunit beta-3 variant, termed beta-3' (see also R93929).
CC      This cytoplasmic domain or its subsequences (R93927-28) are used
CC      to produce anti-integrin antibodies useful in the analysis of
CC      cellular integrin content, in the determination of tumour origin,
CC      and in quantifying the complement of adhesion receptors on cells.
XX
SQ      Sequence 17 AA;
XX
Query Match: 77.8%; Score 28; DB 17; Length 17;
Best Local Similarity 83.3%; Pred. No. 31;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
XX
QY      1 REGAGR 6
DE      | : | | |
DE      2 rdaqr 7
XX
RESULT 4
R93928
ID      R93928 standard; Peptide; 26 AA.
XX
AC      R93928;
XX
DI      14-MAY-1998 (first entry)
XX
DE      Human cytoplasmic domain of the integrin beta3' subunit peptide.
XX
KW      Human; integrin; beta3' subunit; cytoplasmic domain; antibody;
XX      immunoassay.
XX
OS      Homo sapiens.
XX

```

```

DB      6 rdaqr 11
XX
RESULT 3
R93926
ID      R93926 standard; Peptide; 26 AA.
XX
AC      R93926;
XX
DI      20-MAY-1996 (first entry)
XX
DE      Integrin subunit beta-3' cytoplasmic domain (aa4-29).
XX
KW      Integrin subunit beta-3' cytoplasmic domain; antibody; tumour.
XX
OS      Homo sapiens.
XX
PN      J5543664-A.
XX
PD      12-MAR-1996.
XX
PF      25-MAY-1989; 89DS-0357024.
XX
PR      25-MAY-1989; 89DS-0357024.
PR      09-NOV-1992; 92DS-0973547.
PR      10-MAY-1994; 94DS-0240967.
XX
PA      (LJBL) LA JOLLA CANCER RES FOUND.
XX
PI      Roostlahti EJ;
XX
DE      WPI; 1996-159740/16.
XX
PT      Peptide comprising cytoplasmic domain of integrin subunit beta-3'
PT      used to produce anti-integrin antibodies, e.g. for cell integrin
PT      content analysis for the determination tumour origin
XX
PS      Claim 1; Column 7-8; 8pp; English.
XX
CC      A novel cytoplasmic domain sequence (R93926) was identified in an
CC      integrin subunit beta-3 variant, termed beta-3' (see also R93929).
CC      This cytoplasmic domain or its subsequences (R93927-28) are used
CC      to produce anti-integrin antibodies useful in the analysis of
CC      cellular integrin content, in the determination of tumour origin,
CC      and in quantifying the complement of adhesion receptors on cells.
XX
SQ      Sequence 26 AA;
XX
Query Match: 77.8%; Score 28; DB 17; Length 26;
Best Local Similarity 83.3%; Pred. No. 47;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
XX
QY      1 REGAGR 6
DE      | : | | |
DE      15 rdaqr 20
XX
RESULT 4
R93942
ID      R93942 standard; Peptide; 26 AA.
XX
AC      R93942;
XX
DI      14-MAY-1998 (first entry)
XX
DE      Human cytoplasmic domain of the integrin beta3' subunit peptide.
XX
KW      Human; integrin; beta3' subunit; cytoplasmic domain; antibody;
XX      immunoassay.
XX
OS      Homo sapiens.
XX

```

```

FN 08570006 A.
XX
XX 24 JUN 1997.
XX
XX 01 JUN 1999; 990S 0459246.
XX
XX 25 MAY 1999; 990S 0657024.
XX 09 NOV 1997; 990S 0976547.
XX 10 MAY 1997; 990S 0246067.
XX 01 JUN 1997; 990S 0459246.
XX
XX (1997) LA POLLA CANCER RES FOUND.
XX
XX 06570006 E1;
XX
XX WPI: 1998 06243206.
XX
XX Interleukin subunit beta 4' cytoplasmic domain peptide - useful for
XX producing antibodies for interleukin immunassays
XX
XX Claim 1: Cytokine 4'; 7pp; English.
XX
XX The present sequence represents a peptide of the cytoplasmic domain
XX of interleukin subunit beta 4'. The peptide can be used for producing
XX antibodies for use in immunassays for interleukins.
XX
XX Sequence 26 AA:

Query Match 77.8% Score 28; DB 19; Length 26;
Best Local Similarity 83.0% Prod. No. 47;
Matches 5; Conservation 1; Mismatches 0; Indels 0; Gaps 0;

25 1 REGACK 6
DB 18 rdquad 24
1:1111
18 rdquad 24

RESULT 6
Y44473
ID Y44473 standard; peptide; 9 AA.
XX
XX Y44473;
XX
XX 27-MAR-2000 (first entry)
DI
DI Human antithrombin 111 variant 14.A (residues 485-493).
DE
XX
XX Human; antithrombin 111; AT111 variant 14.A; elastase resistant;
XX 1q5 activated neutrophil resistant; anti thrombin activity; heparin;
XX anti factor Xa activity; blood clotting disorder; sepsis; trauma; stroke;
XX thrombin activation-related pathological symptom; restenosis; thrombosis;
XX acute respiratory distress syndrome; thromboembolism; recanalisation.
XX
XX Homo sapiens.
XX Synthetic.
XX
XX Key location/qualifiers
XX
XX FT Mismatch 5
XX
XX FT Mismatch 6
XX
XX FT Mismatch 6
XX
XX W09958098-A2.
XX
XX 18 NOV 1999.
XX
XX 12 MAY 1999; 990S 0810549.
XX
XX 12 MAY 1998; 990S 0085197.
XX 05-MAY-1999; 990S-0085197.
XX
XX (5-767) PCR S.C.
XX (PICAV) PICARD V.
XX (ZENE) ZENEHUB P
XX
XX Bock SC, Picard V, Zendeck-Rohr F;
XX
XX WPI: 2000-116274/10
XX
XX New modified human antithrombin 111 compounds, used for treating e.g.
XX sepsis, trauma, acute respiratory distress syndrome, restenosis,
XX thrombosis, thromboembolism or stroke
XX
XX Claim 13: Page 57; 7pp; English.
XX
XX The present sequence is from an antithrombin 111 (AT111) variant, 14.A

```

```

FN 08570006 A.
XX
XX 24 JUN 1997.
XX
XX 01 JUN 1999; 990S 0459246.
XX
XX 25 MAY 1999; 990S 0657024.
XX 09 NOV 1997; 990S 0976547.
XX 10 MAY 1997; 990S 0246067.
XX 01 JUN 1997; 990S 0459246.
XX
XX (1997) LA POLLA CANCER RES FOUND.
XX
XX 06570006 E1;
XX
XX WPI: 1998 06243206.
XX
XX Interleukin subunit beta 4' cytoplasmic domain peptide - useful for
XX producing antibodies for interleukin immunassays
XX
XX Claim 1: Cytokine 4'; 7pp; English.
XX
XX The present sequence represents a peptide of the cytoplasmic domain
XX of interleukin subunit beta 4'. The peptide can be used for producing
XX antibodies for use in immunassays for interleukins.
XX
XX Sequence 26 AA:

Query Match 77.8% Score 28; DB 19; Length 26;
Best Local Similarity 83.0% Prod. No. 47;
Matches 5; Conservation 1; Mismatches 0; Indels 0; Gaps 0;

25 1 REGACK 6
DB 18 rdquad 26
1:1111
18 rdquad 26

RESULT 6
Y44473
ID Y44473 standard; peptide; 26 AA.
XX
XX Y44473.
XX
XX 20 MAY 1999 (first entry)
DI
DI Interleukin subunit beta 4' cytoplasmic domain c-terminal sequence.
XX
XX Interleukin subunit beta 4' cytoplasmic domain; antibody; tumour.
XX
XX Homo sapiens.
XX
XX 08570006 A.
XX
XX 12 MAR 1999.
XX
XX 25 MAY 1999; 990S 0657024.
XX
XX 25 MAY 1999; 990S 0657024.
XX 09 NOV 1997; 990S 0976547.
XX 10 MAY 1997; 990S 0246067.
XX
XX (1997) LA POLLA CANCER RES FOUND.
XX
XX 06570006 E1;
XX
XX WPI: 1998 159746/16
XX N 8506; 116856.
XX
XX Peptide comprising cytoplasmic domain of interleukin subunit beta-4'
XX used to produce anti interleukin antibodies, e.g. for cell interleukin
XX content analysis for the determination tumour origin
XX

```

CC derived from human ATIII.N135A cDNA insert of the pBlueBac baculovirus
 CC expression construct and comprises residues 365-383. The variant has
 CC improved resistance to elastase and 1q3-activated neutrophils while
 CC retains anti-thrombin and anti-factor Xa activities. It may be
 CC expressed as glycoforms with enhanced heparin affinity which target the
 CC blood vessel wall more efficiently than ATIII with normal heparin
 CC affinity. The modified ATIII can be used to treat thrombin activation-
 CC related pathological symptoms due to sepsis, trauma, acute
 CC respiratory distress syndrome, restenosis, thrombosis, thromboembolism
 CC and stroke. It can also be used to reduce the risk of reocclusion
 CC and restenosis in percutaneous transluminal coronary angioplasty,
 CC thrombolysis associated with surgery, ischaemia/reperfusion injury, and
 CC coagulation abnormalities in cancer or surgical patients.
 XX
 SQ Sequence: 9 AA:

 Query Match: 72.2%; Score 26; DB 21; Length 9;
 Best Local Similarity: 100.0%; Pred. No. 3.2e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

 QY 2 EGACR 6
 Db 11111
 5 exact 9

 RESULT 7
 Y44474
 ID Y44474 standard; peptide: 9 AA.
 AC Y44474:
 DT 27-MAR-2000 (first entry)
 DE Human antithrombin III variant 13.B (residues 385-393).
 KW Human: antithrombin III; ATIII variant 13.B; elastase-resistant;
 KW 1q3-activated neutrophil resistant; anti-thrombin activity; heparin;
 KW anti-activated Xa activity; blood clotting disorder; sepsis; trauma; stroke;
 KW thrombin activation-related pathological symptom; restenosis; thrombosis;
 KW acute respiratory distress syndrome; thromboembolism; reocclusion.
 OS Homo sapiens.
 SC Synthetic.
 EH Key: Location/Qualifiers
 FT Misc-difference 4 /note "ATIII.N135A Ala at 387 is substituted by Gly"
 FT Misc-difference 5 /note "ATIII.N135A Val at 389 is substituted by Gly"
 FT Misc-difference 6 /note "ATIII.N135A Ile at 390 is substituted by Gly"
 XX
 IN W0958098-A2.
 XX 18-NOV-1999.
 XX 12-MAY-1999; 99NO-US10549.
 XX 12-MAY-1998; 98US-0085197.
 XX 05-MAY-1999; 99US-0085197.
 XX (BWF7) ROCK S C.
 XX (PICAV) PICARD V.
 XX (ZENG) ZENGJIEHUI F.
 XX Rock SC, Picard V, Zengjiehui F;
 XX WPI: 2000 116274/36
 XX New modified human antithrombin III compounds, used for treatment of:
 XX sepsis, trauma, acute respiratory distress syndrome, restenosis,
 XX thrombosis, thromboembolism or stroke.

XX
 PS Claim 13; Page 57; 75pp; English.
 CC The present sequence is from an antithrombin III (ATIII) variant, 13.B,
 CC derived from human ATIII.N135A cDNA insert of the pBlueBac baculovirus
 CC expression construct and comprises residues 385-393. The variant has
 CC improved resistance to elastase and 1q3-activated neutrophils while
 CC retains anti-thrombin and anti factor Xa activities. It may be
 CC expressed as glycoforms with enhanced heparin affinity which target the
 CC blood vessel wall more efficiently than ATIII with normal heparin
 CC affinity. The modified ATIII can be used to treat thrombin activation-
 CC related pathological symptoms due to sepsis, trauma, acute
 CC respiratory distress syndrome, restenosis, thrombosis, thromboembolism
 CC and stroke. It can also be used to reduce the risk of reocclusion
 CC and restenosis in percutaneous transluminal coronary angioplasty,
 CC thrombolysis associated with surgery, ischaemia/reperfusion injury, and
 CC coagulation abnormalities in cancer or surgical patients.
 XX
 SQ Sequence: 9 AA:

 Query Match: 72.2%; Score 26; DB 21; Length 9;
 Best Local Similarity: 100.0%; Pred. No. 3.2e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

 QY 2 EGACR 6
 Db 11111
 5 exact 9

 RESULT 8
 Y44475
 ID Y44475 standard; peptide: 9 AA.
 AC Y44475:
 DT 27-MAR-2000 (first entry)
 DE Human antithrombin III variant 13.C (residues 385-394).
 KW Human: antithrombin III; ATIII variant 13.C; elastase-resistant;
 KW 1q3-activated neutrophil resistant; anti-thrombin activity; heparin;
 KW anti-factor Xa activity; blood clotting disorder; sepsis; trauma; stroke;
 KW thrombin activation-related pathological symptom; restenosis; thrombosis;
 KW acute respiratory distress syndrome; thromboembolism; reocclusion.
 OS Homo sapiens.
 SC Synthetic.
 EH Key: Location/Qualifiers
 FT Misc-difference 4 /note "ATIII.N135A Ala at 387 is substituted by Gly"
 FT Misc-difference 5 /note "ATIII.N135A Val at 388 is substituted by Leu"
 FT Misc-difference 6 /note "ATIII.N135A Val at 389 is substituted by Gly"
 FT Misc-difference 6 /note "ATIII.N135A Ile at 390 is substituted by Gly"
 XX
 IN W0958098-A2.
 XX 18-NOV-1999.
 XX 12-MAY-1999; 99NO-US10549.
 XX 12-MAY-1998; 98US-0085197.
 XX 05-MAY-1999; 99US-0085197.
 XX (BWF7) ROCK S C.
 XX (PICAV) PICARD V.
 XX (ZENG) ZENGJIEHUI F.
 XX Rock SC, Picard V, Zengjiehui F;
 XX WPI: 2000 116274/36


```

PR 18-DEC-1997; 97US-0070923.
PR 19-DEC-1997; 97US-0068169.
PR 19-DEC-1997; 97US-0068365.
PR 19-DEC-1997; 97US-0068367.
PR 19-DEC-1997; 97US-0068368.
XX (HUMAN) HUMAN GENE: SC1 INC.
XX
PI Carter KC, Duan RD, Feng P, Ferric AM, Florence C;
PI Florence K, Greene JM, Janat F, Kyaw H, Moore PA;
PI Ni J, Rosen CA, Ruben SM, Shi Y, Soppet DR, Wei Y;
PI Yu G;
XX WPI: 1999-418749/35
XX
XX New isolated human genes encoding secreted polypeptides
XX
XX Disclosure: Page 460; 537pp; English.
XX
XX X47916 to X48024 represent 110 isolated human secreted protein genes.
XX X47224 to X47627 represent the secreted proteins encoded by the 110
XX human genes. The genes and their corresponding secreted polypeptides are
XX useful for preventing, treating or ameliorating medical conditions,
XX e.g. by protein or gene therapy. Also pathological conditions can be
XX diagnosed by determining the amount of the new polypeptides in a sample
XX or by determining the presence of mutations in the new genes. Specific
XX genes are described for each of the 110 genes, based on which tissues they
XX are most highly expressed in, and include developing products for the
XX diagnosis or treatment of cancer, tumors, developmental abnormalities
XX and total deficiencies, blood disorders, diseases of the immune system,
XX autoimmune diseases, inflammation, allergies, Alzheimer's and cognitive
XX disorders, schizophrenia, arthritis, asthma, psoriasis, sepsis, skin
XX disorders, atherosclerosis, diabetes, cardiovascular disorders, kidney
XX disorders, digestive disorders, fibrosis, infections and AIDS. The
XX polypeptides are also useful for identifying their binding partners.
XX The sequences given in X47916 to X47919 and X47921 are used in the
XX exemplification of the present invention.
XX
XX Sequence 26 AA:
SQ
Query Match 72.2%; Score 26; DB 29; Length 26;
Best Local Similarity 93.9%; Pred. NO. 1.1e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 RESAGR 6
DB 10 readmr 15
III II
XX
RESULT 11
Y52526
ID Y52526 standard; peptide; 25 AA.
XX
AC Y52526;
XX
DE 22-FEB-2000 (first entry)
XX
DE House dust mite allergen 60 kD protein (mapd) N terminal fragment.
XX
XX Mite allergen protein; map: high molecular weight; HMW-map; allergy;
XX house dust mite; IgE; immunoglobulin E; allergen; mapd;
XX hypersensitivity reaction; therapy; treatment; diagnosis; human, feline;
XX canine; veterinary; antibody; vaccine; immunization.
XX
OS Dermatophagoides farinae.
XX
XX Key Location/Qualifiers
XX Misc-difference 1
XX /label- Xaa
XX /note- "Xaa - any amino acid"
XX
XX W0954449-A2.

```

```

XX 28-OCT-1999.
XX
XX 16-APR-1999; 99WO-US08524.
XX
XX 17-APR-1998; 98US-0062013.
XX 14 MAY 1998; 98US 0085295.
XX 02-SEP-1998; 98US-0098909.
XX (HESK-) HESKA CORP.
XX
XX McCall CA, Hunter SW, Weber ER;
XX WPI: 2000-052700/04.
XX
XX Novel high molecular weight Dermatophagoides mite allergen acid polypeptides
XX used to modify an animal's hypersensitivity to mite allergens.
XX
XX Claim 3, Page 81, 154pp; English.
XX
XX This sequence represents an N-terminal fragment of the Dermatophagoides
XX farinae 60 kD mite allergen protein (mapd). The 60 kD allergen was
XX isolated from a 3.7 kilobase fragment by gel filtration. With each
XX fraction being analysed for the presence of proteins that bound to IgE
XX present in allergic dog antisera. Mite allergenic proteins and
XX peptides, and nucleic acids encoding them, may be used in therapeutic
XX compositions to modify an animal's hypersensitivity reaction to mite
XX allergens. Animals that may be treated include mammals and birds,
XX especially felines, canines, equines, humans, other pets, and work or
XX domestic animals. The proteins or fragments may also be used to
XX diagnose allergies via a skin test. The proteins and peptides can also
XX be used to raise antibodies, which have a variety of potential
XX uses. For example, they can be used as vaccines to passively immunize
XX animals against test mite hypersensitivity, as positive controls in
XX test kits and as tools to recover desired dust mite allergens from a
XX mixture of proteins.
XX
XX Sequence 25 AA:
SQ
Query Match 69.4%; Score 25; DB 21; Length 25;
Best Local Similarity 57.1%; Pred. NO. 1.7e+02;
Matches 1; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 1 PEGAGPM 7
DB 17 rqaqkm 23
III III
XX
RESULT 12
R32448
ID R32448 standard; peptide; 10 AA.
XX
AC R32448;
XX
DE 10-JUN-1993 (first entry)
XX
DE Human thrombospondin type I repeat derived peptide #21.
XX
XX thrombosis; coagulation; heparin binding inhibitor; type I repeat.
XX
OS Synthetic.
XX
CN U57801812-A.
XX
PD 15-DEC-1992.
XX
XX 06-DEC-1991; 91US 0801812.
XX
XX 06-DEC-1991; 91US-0801812.
XX
XX (USSH ) US DEPT HEALTH & HUMAN SERVICE.
XX

```

P1 Roberts lab;
 XX WP1: 1995-06749/mg.
 DR
 XX New sulphated glycoconjugate binding peptide(s) - from type 1
 P1 repeats of human thrombospondin, preventing interaction of the
 P1 glycoconjugates with adhesion molecules, growth factors, etc.
 XX
 PS Disclosure: Page 13; 64pp; English.
 XX
 CC This peptide was obtained from the adhesive glycoprotein
 CC thrombospondin. It was tested for inhibitory activity against
 CC binding to heparin and laminin and was inactive. The preferred
 CC peptides of the invention (which do have inhibitory activity) all
 CC contain a subsequence WSXW (X: D, E, G, A, S) with a substantial
 CC lack of an electrical charge. They act as inhibitors of
 CC heparin or related sulphated glycoconjugate-binding to adhesion
 CC molecules, growth factors, etc.
 XX
 SQ Sequence: 10 AA;
 Query Match 66.7%; Score 24; DB 14; Length 10;
 Best Local Similarity 66.7%; Pred. No. 1,1e+02;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 QY 1 RQZAKR 6
 DB 1 RKAQR 6
 RESULT 14
 R79366
 ID R79366 standard; peptide; 19 AA
 XX AC
 XX R79366;
 DT 13-OCT-1995 (first entry)
 XX
 DE Thrombospondin heparin binding peptide Hept.
 XX
 KW Thrombospondin: inhibitor; type 1 repeat unit; identification;
 KW extracellular matrix protein; heparin; proliferation; adhesion;
 KW met 1117; extravasation; neovascularisation; metastatic tumour;
 KW breast carcinoma; melanoma; cytotatatic; heparin sulphate; R3F 2;
 KW fibroblast growth factor 2; sulphate glycoconjugate; Kaposi's Sarcoma,
 KW haemangioma; diabetic retinopathy.
 XX
 OS Homo Sapiens.
 CS Synthesized.
 XX
 PN US6051549 A
 ID 18 APR 2000
 XX
 XX 11 MAR 1999; 9605 0941119.
 XX
 XX 07 JUN 1995; 9605-0487568.
 XX 06 DEC 1991; 9105 0801812.
 XX 21 MAR 1994; 9405-0215085.
 XX
 (US6051549) US DEPT HEALTH & HUMAN SERVICES.
 XX
 XX Chao N, Krut zsch HC, Roberts DG;
 XX WP1: 1995-44947/29.
 DR
 XX New polypeptides binding to heparin or related sulfated glycoconjugates
 P1 are useful for inhibiting endothelial cell proliferation and tumor
 P1 growth.
 XX
 XX Example 1; Column 18; 13pp; English.
 PS
 XX

CC The present invention describes a polypeptide (1) binding to heparin or
 CC related sulphated glycoconjugates with high affinity. R62638 to R62747
 CC represent peptides derived from human thrombospondin, which are used in
 CC the exemplification of the present invention. (1) can be used for
 CC inhibiting heparin or heparin sulphate interaction in a sample, without
 CC activating latent transforming growth factor (TGF)-beta, for inhibiting
 CC interaction of heparin or heparin sulphate with R3F-2 in a sample, and
 CC for inhibiting endothelial cell proliferation, and tumour growth, in
 CC patients. The peptides bind heparin or related sulphate glycoconjugates
 CC with high affinity. The peptides or their conjugates can be used in
 CC blocking or modifying the action on cellular processes of heparin (e.g.,
 CC proliferation, adhesion, motility, extracellular matrix remodelling),
 CC sulphatides, related sulphated glycoconjugates, fibronectin, and basic
 CC fibroblast growth factor, involving malignant cell lines and normal
 CC endothelial cells. The peptides can also be used in the treatment of
 CC metastatic tumours, breast carcinomas, melanomas, Kaposi's sarcomas,
 CC haemangiomas, diabetic retinopathies, and various pathological
 CC conditions dependent on neovascularisation.
 XX
 SQ Sequence: 10 AA;
 Query Match 66.7%; Score 24; DB 21; Length 10;
 Best Local Similarity 66.7%; Pred. No. 1,1e+02;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 QY 1 RQZAKR 6
 DB 1 RKAQR 6
 RESULT 14
 R79366
 ID R79366 standard; peptide; 19 AA
 XX AC
 XX R79366;
 DT 13-OCT-1995 (first entry)
 XX
 DE Thrombospondin heparin binding peptide Hept.
 XX
 KW Thrombospondin type 1 repeat sequence; transforming growth factor beta;
 KW wound healing; fibrosis; endothelial cell proliferation; heparin;
 XX Synthetic.
 CS
 XX W09505191-A.
 PN
 XX 23-FEB-1995.
 PD
 XX 12-AUG-1994; 94W0-0509194
 XX
 XX 14-AUG-1994; 94US-0106120.
 PR 04-MAY-1994; 94US-0238169.
 XX
 XX (DAHP-) DAB RES FOUND.
 PA
 XX Krut zsch HC, Murphy-Blirrich JE, Roberts DG, Schellz Cherry S;
 P1 WP1: 1995-098579/13.
 DR
 XX Stimulating or inhibiting transforming growth factor beta by
 P1 contacting with thrombo-spondin or an activating enzyme used
 P1 to enhance wound healing or prevent fibrosis
 XX
 XX Examples, Page 28, 67pp, English.
 PS
 XX The sequence of a peptide lacking the thrombospondin heparin binding
 CC consensus motif: Trp-Ser-His-Tyr. The peptide corresponds to amino acid
 CC 17-35 of thrombospondin. The peptide was tested for the ability to
 CC stimulate the conversion of TGF beta from the latent to active form.
 CC Peptides (see R6766-79) which stimulate the conversion of latent TGF-beta
 CC to active TGF-beta can be used to enhance wound healing whilst inhibitory

CC peptides (R69780-90) can be used to prevent fibrosis or block TGF- β
 mediated endothelial cell proliferation.

SO Sequence 19 AA;

Query Match 66.7%; Score 24; DB 16; Length 19;

Best Local Similarity 66.7%; Posed No. 2002;

Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 REGAR 6

Db 1:1:1

7 rkasqr 12

RESULT 15

R97641

10 R97641 standard; peptide; 19 AA.

XX AC

XX R97641;

XX AC

DE 04-FEB-1997 (first entry)

XX DE Cell binding domain peptide HRD I of thrombospondin 1.

XX Thrombospondin 1; cell-binding domain; multidomain glycoprotein;

XX cell migration; cell attachment; cell differentiation; protein receptor;

XX integral membrane glycoprotein; cell adhesion; inflammation; arthritis;

XX angiogenesis; cancer metastasis.

XX Synthetic.

OS

XX W 9617059-A2.

XX FN

XX 06-JUN-1996.

XX FN

XX 21-NOV-1995; 95WO-0514553.

XX XX

XX 30-NOV-1994; 94US-0347000.

XX XX

XX (UNIV) UNIV WASHINGTON.

XX PA

XX Frazier WA, Gao A;

XX XX

XX WBI: 1986-277779/28.

XX XX

PI Isolated thrombospondin 1 receptor protein - useful for studying

PI processes involved in inflammation, angiogenesis, arthritis and

PI cancer metastasis

XX XX

XX Example: Fig 1; 43pp; English.

PS

XX R97634-R97644 represent cell binding domain (CBD) peptides of

CC thrombospondin 1 (TSP1). Thrombospondins are a family of multidomain

CC glycoproteins that influence migration, attachment, and differentiation

CC of cell types. These CBD peptide sequences were used to identify the

CC protein of the invention. The protein of the invention is a 52 kD

CC protein receptor of the CBD peptides of TSP1. The receptor protein is an

CC integral membrane glycoprotein. The binding site of the protein receptor

CC is accessible on the intact cell surface, and the receptor mediates

CC adhesion of cells to the CBD of TSP1. Prior blockade of the protein

CC receptor by reaction of cells with the CBD peptides and a crosslinking

CC reagent inhibits cell adhesion. The receptor protein can be used to

CC study the migration, attachment, proliferation and differentiation of a

CC number of cell types. The receptor can also be used for assessing the

CC processes of inflammation, angiogenesis, arthritis and cancer metastasis.

XX

SO Sequence 19 AA;

Query Match

Best Local Similarity 66.7%; Score 24; DB 17; Length 19;

Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

•Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 REGAR 6

Db 1:1:1

7 rkasqr 12

Search completed: May 1, 2001, 11:25:31

Job time: 101 sec




```
Best Local Similarity 82.3%; Pred. No. 9;
Matches 5; Conservative 1; Mismatches 0; Gaps 0;

QY 1 REGAGR 6
   1:||||
DB 6 RDGAGR 11

RESULT 3
5498694-9
: Patent No. 5498694
: APPLICANT: ROGSLAHTI, ERKKI I.
: TITLE OF INVENTION: PEPTIDES OF THE CYTOPLASMIC DOMAIN OF
: INTEGRIN
: NUMBER OF SEQUENCES: 12
: CURRENT APPLICATION DATA:
:   APPLICATION NUMBER: 93-08-241,567
:   FILING DATE: 10-MAY-1994
: PRIOR APPLICATION DATA:
:   APPLICATION NUMBER: 973,547
:   FILING DATE: 09-NOV-1992
:   APPLICATION NUMBER: 357,824
:   FILING DATE: 25-MAY-1989
: SEQ ID NO:4
:   LENGTH: 26
5498694-9

Query Match 77.8%; Score 28; DB 6; Length 26;
Best Local Similarity 83.3%; Pred. No. 14;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 REGAGR 6
   1:||||
DB 15 RDGAGR 20

RESULT 4
5498694-4
: Patent No. 5498694
: APPLICANT: ROGSLAHTI, ERKKI I.
: TITLE OF INVENTION: PEPTIDES OF THE CYTOPLASMIC DOMAIN OF
: INTEGRIN
: NUMBER OF SEQUENCES: 12
: CURRENT APPLICATION DATA:
:   APPLICATION NUMBER: 93-08-241,567
:   FILING DATE: 10-MAY-1994
: PRIOR APPLICATION DATA:
:   APPLICATION NUMBER: 973,547
:   FILING DATE: 09-NOV-1992
:   APPLICATION NUMBER: 357,824
:   FILING DATE: 25-MAY-1989
: SEQ ID NO:4
:   LENGTH: 29
5498694-4

Query Match 77.8%; Score 28; DB 6; Length 29;
Best Local Similarity 83.3%; Pred. No. 15;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 REGAGR 6
   1:||||
DB 18 RDGAGR 23

RESULT 5
US-08-279-754-5
: Sequence 5, Application US-08-279-754
: Patent No. 5486470
: GENERAL INFORMATION:
: APPLICANT: DARRK, PAUL L.
: APPLICANT: HALL, DAWN L.
```

```
APPLICANT: KUC, LAWRENCE C.
TITLE OF INVENTION: PURIFIED HERPES SIMPLEX VIRAL PROTHASE
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: JOANNE M. GIESSE
STREET: P.O. BOX 2000, 126 E. LINCOLN AVE.
CITY: RAHWAY
STATE: NJ
COUNTRY: USA
ZIP: 07065
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US-08-279-754
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: GIESSE, JOANNE M.
REGISTRATION NUMBER: 32,838
REFERENCE/BACKLIT NUMBER: 19233
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908) 594-3046
TELEFAX: (908) 594-4720
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 14 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
FRAGMENT TYPE: Internal
US-08-279-754-5

Query Match 69.4%; Score 25; DB 1; Length 14;
Best Local Similarity 71.4%; Pred. No. 29;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 REGAGR 7
   1:||||
DB 8 REGAGR 14

RESULT 6
US-08-687-820-4
: Sequence 4, Application US-08-687-820
: Patent No. 5972666
: GENERAL INFORMATION:
: APPLICANT: Hippenmeyer, Paul J
: APPLICANT: Luckow, Verne A
: APPLICANT: Rankin, Anne M
: TITLE OF INVENTION: Assembly Deficient Herpesvirus Vaccine
: NUMBER OF SEQUENCES: 4
: CORRESPONDENCE ADDRESS:
ADDRESSEE: Joseph W. Bullock, G.E. Seale & Co.,
ADDRESSEE: Corporate Patent Dept.
STREET: P.O. Box 5110
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60680
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.40
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US-08-687-820
FILING DATE: 26-MAY-1996
```



```

1 CLASSIFICATION: 4.4
2 ATTORNEY/AGENT INFORMATION:
3 NAME: Bullock, Joseph W
4 REGISTRATION NUMBER: 37,193
5 REFERENCE NUMBER: C 2477
6 TELECOMMUNICATION INFORMATION:
7 TELEPHONE: 414.694.9094
8 TELEFAX: 414.694.9095
9 INFORMATION FOR SEQ ID NO: 4:
10 SEQUENCE CHARACTERISTICS:
11 LENGTH: 14 amino acids
12 TYPE: amino acid
13 SUPPLEMENTARY DATA:
14 Synology: Linear
15 MOLECULE TYPE:
16 DESCRIPTION: 23.3 "CYTRETIN"
17 US 09-341-829A-5

```

```

Query Match: 69.44% Score 25; ID 5; Length 14;
Best Local Similarity: 67.14% Prod. No. 29;
Matches: 4; Conservative: 4; Mismatches: 0; Gaps: 0;

```

```

57 1 REGARM 7
58 11111
59 1 REGARM 7

```

```

1 RESULT 7
2 US 09-341-829A-5
3 Sequence 5, Application 10/US9609652
4 GENERAL INFORMATION:
5 APPLICANT: LARKE, PAUL G.
6 APPLICANT: HALL, DAWN L.
7 APPLICANT: KEE, LAWRENCE C.
8 TITLE OF INVENTION: PERTINENT HERPES SIMPLEX VIRAL PROLIFERATION
9 NUMBER OF SEQUENCES: 5
10 CURRENT APPLICATION DATA:
11 ADDRESSER: JAMES M. GLESSER
12 STREET: P.O. Box 2000, 125 E. LINCOLN AVE.
13 CITY: FAIRWAY
14 STATE: NJ
15 COUNTRY: USA
16 ZIP: 07410
17 COMPUTER READABLE FORM:
18 MEDIUM TYPE: Floppy disk
19 COMPUTER: IBM PC compatible
20 OPERATING SYSTEM: PC-DOS/MS-DOS
21 SOFTWARE: Patent In Release #1.0, Version #1.25
22 CURRENT APPLICATION DATA:
23 APPLICATION NUMBER: 10/US9609652
24 FILING DATE:
25 CLASSIFICATION:
26 ATTORNEY/AGENT INFORMATION:
27 NAME: GLESSER, JAMES M.
28 REGISTRATION NUMBER: 42,849
29 REFERENCE NUMBER: 19234
30 TELECOMMUNICATION INFORMATION:
31 TELEPHONE: (908) 594.3046
32 TELEFAX: (908) 594.4720
33 INFORMATION FOR SEQ ID NO: 5:
34 SEQUENCE CHARACTERISTICS:
35 LENGTH: 14 amino acids
36 TYPE: amino acid
37 SUPPLEMENTARY DATA:
38 Synology: Linear
39 MOLECULE TYPE: peptide
40 HYDROPHOBICITY: No
41 EXAMINATION TYPE: Internal
42 US 09-341-829A-5

```

```

Query Match: 69.44% Score 25; ID 5; Length 14;

```

```

1 Best Local Similarity: 71.44% Prod. No. 29;
2 Matches: 5; Conservative: 1; Mismatches: 1; Gaps: 0;
3
4 57 1 REGARM 7
5 11111
6 8 REGARM 14
7
8 RESULT 8
9 US-07-801-812A-21
10 Sequence 21, Application US/07801812A
11 Patent No. 5357041
12 GENERAL INFORMATION:
13 APPLICANT: David D. Roberts et al
14 TITLE OF INVENTION: HEPATITIS AND SPATIALLY-BINDING
15 TITLE OF INVENTION: PEPTIDES FROM THE TYPE 1 REPEATS OF HEMAN
16 TITLE OF INVENTION: THROMBOSPONDIN FROM THE MELANOMA CELL ADHESION
17 NUMBER OF SEQUENCES: 27
18 CORRESPONDENCE ADDRESS:
19 ADDRESSER: Lowe, Peter, Ladiges & Becker
20 STREET: Suite 300, 99 Canal Center Plaza
21 CITY: Alexandria
22 STATE: Virginia
23 COUNTRY: USA
24 ZIP: 22314
25 COMPUTER READABLE FORM:
26 MEDIUM TYPE: Floppy disk
27 COMPUTER: IBM PC compatible
28 OPERATING SYSTEM: PC-DOS/MS-DOS
29 SOFTWARE: DOS TEXT FILE
30 CURRENT APPLICATION DATA:
31 APPLICATION NUMBER: US/07801812A
32 FILING DATE: 19911206
33 CLASSIFICATION: 530
34 ATTORNEY/AGENT INFORMATION:
35 NAME: J.G. Mullins
36 REGISTRATION NUMBER: 33,074
37 REFERENCE NUMBER: 217,111
38 TELECOMMUNICATION INFORMATION:
39 TELEPHONE: 703.684.1111
40 INFORMATION FOR SEQ ID NO: 21:
41 SEQUENCE CHARACTERISTICS:
42 LENGTH: 10 amino acids
43 TYPE: AMINO ACID
44 TOPOLOGY: Linear
45 MOLECULE TYPE: peptide
46 US-07-801-812A-21
47
48 Query Match: 66.76% Score 24; ID 1; Length 10;
49 Best Local Similarity: 66.76% Prod. No. 33;
50 Matches: 4; Conservative: 2; Mismatches: 0; Gaps: 0;
51
52 57 1 REGARM 6
53 11111
54 1 REGARM 6
55
56 RESULT 9
57 US 08-487-568-21
58 Sequence 21, Application US/08487568
59 Patent No. 6770564
60 GENERAL INFORMATION:
61 APPLICANT: Roberts, David D.
62 APPLICANT: Brown, Philip J.
63 APPLICANT: Bryant, Joseph L.
64 APPLICANT: Imman, John K.
65 APPLICANT: Krutzsch, Henry C.
66 APPLICANT: Guo, Nenghua
67 TITLE OF INVENTION: Heparin and Sulfate Binding Peptides
68 TITLE OF INVENTION: From the Type 1 Repeats of Human Trombospondin and
69 TITLE OF INVENTION: Conjugates Thereof
70 NUMBER OF SEQUENCES: 114

```

```

CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew
STREET: 200 Market Plaza, Steuart Street Tower
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94105-1492
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.10
CURRENT APPLICATION DATA:
FILING DATE: 07-JUN-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/215,085
FILING DATE: 21-MAR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07,991,812
FILING DATE: 06-DEC-1991
ATTORNEY/AGENT INFORMATION:
NAME: Dow, Kuroki B.
REGISTRATION NUMBER: 29,684
REFERENCE NUMBER: 015290-023310
TELEPHONE: (415) 326-2400
TELEFAX: (415) 326-2422
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 amino acids
TYPE: amino acid
STRANGLINNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-341-829a-5

```

Query Match: 66.7% Score 24; DB 1; Length 10;

Best Local Similarity 66.7% Pred. No. 34;

Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 PEGAGR 6
   1 111

```

```

DB 1 PEGAGR 6

```

RESULT 10

```

US-08-347-000-10
Sequence 10, Application US/08447000
Patent No. 5622665
GENERAL INFORMATION:
APPLICANT: Frazier, William A.
APPLICANT: Frazier, William A.
TITLE OF INVENTION: Receptor for Cell-Binding Domain of
TITLE OF INVENTION: Thrombospondin
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Scott J. Meyer, Monsanto/Scarle, A8SG
STREET: 800 N. Lindbergh Blvd.
CITY: St. Louis
STATE: Missouri
COUNTRY: USA
ZIP: 63167
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.10
CURRENT APPLICATION DATA:
FILING DATE: 03-MAR-1994

```

```

CLASSIFICATION: 540
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/029,434
FILING DATE: 05-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: Meyer, Scott J.
REGISTRATION NUMBER: 25,275
REFERENCE/DOCKET NUMBER: WU-284H
TELECOMMUNICATION INFORMATION:
TELEPHONE: (314) 694-4117
TELEFAX: (314) 694-5435
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 19 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-447-000-10

Query Match: 66.7% Score 24; DB 1; Length 19;
Best Local Similarity 66.7% Pred. No. 64;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 PEGAGR 6
   1 111
DB 7 PEGAGR 12

RESULT 11
US-08-253-171-1
Sequence 11, Application US/08253171
Patent No. 5472863
GENERAL INFORMATION:
APPLICANT: MARGIA, Kazuhiko
APPLICANT: KUBOTA, Michio
APPLICANT: SUGIMOTO, Toshiyuki
APPLICANT: MIYAKE, Toshio
TITLE OF INVENTION: TROPALONE-RELEASING ENZYME, AND ITS
TITLE OF INVENTION: PREPARATION AND USES
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROWNY AND NEIMARK
STREET: 419 Seventh Street, N.W., Suite 300
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.10
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08,253,171
FILING DATE: 02-JUN-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 156338/1994
FILING DATE: 03-JUN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 340343/1994
FILING DATE: 09-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 79291/1994
FILING DATE: 28-MAR-1994
ATTORNEY/AGENT INFORMATION:
NAME: NEIMARK, Sheridan
REGISTRATION NUMBER: 20,520
REFERENCE/DOCKET NUMBER: MARGITA 2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202 629-5197
TELEFAX: 202 747 4528

```

```

1  TELEX: 248633
2  INFORMATION FOR SEQ ID NO: 1:
3  SEQUENCE CHARACTERISTICS:
4  LENGTH: 20 amino acids
5  TYPE: amino acid
6  STRANDEDNESS: single
7  TOPOLOGY: linear
8  MOLECULE TYPE: peptide
9  US 09 341 829
10
11 Query Match
12 Best Local Similarity: 63.9%; Score 23; DB 1; Length 10;
13 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
14
15 QY 2 EAGR 6
16 :||||
17 DB 5 QGAGR 9
18
19 RESULT 14
20 US 09 341 829-1
21 : Sequence 1, Application US/0844352
22 : Patent No. 5591612
23 : GENERAL INFORMATION:
24 : APPLICANT: MARUTA, Kazuhiko
25 : APPLICANT: KUROTA, Michio
26 : APPLICANT: SUGIMOTO, Toshiyuki
27 : APPLICANT: MIYAKE, Toshio
28 : TITLE OF INVENTION: TRIHALOGENE-RELEASING ENZYME, AND ITS
29 : TITLE OF INVENTION: PREPARATION AND USES
30 : NUMBER OF SEQUENCES: 11
31 : CORRESPONDENCE ADDRESS:
32 : ADDRESSEE: BROWDY AND NEWMARK
33 : STREET: 419 Seventh Street, N.W., Suite 400
34 : CITY: Washington
35 : STATE: D.C.
36 : COUNTRY: USA
37 : ZIP: 20004
38 : COMPUTER READABLE FORM:
39 : MEDIUM TYPE: Floppy disk
40 : COMPUTER: IBM PC compatible
41 : OPERATING SYSTEM: PC DOS/MS DOS
42 : SOFTWARE: Patent In Release #1.0, Version #1.25
43 : CURRENT APPLICATION DATA:
44 : APPLICATION NUMBER: US/08/443,352
45 : FILING DATE: 17-MAY-1995
46 : CLASSIFICATION: 435
47 : PRIOR APPLICATION DATA:
48 : APPLICATION NUMBER: US 08/253,171
49 : FILING DATE: 02 JUN-1994
50 : APPLICATION NUMBER: JP 15638/1994
51 : FILING DATE: 24 JUN-1994
52 : PRIOR APPLICATION DATA:
53 : APPLICATION NUMBER: JP 43343/1994
54 : FILING DATE: 09 DEC-1993
55 : PRIOR APPLICATION DATA:
56 : APPLICATION NUMBER: JP 79291/1994
57 : FILING DATE: 28 MAR-1994
58 : ATTORNEY/AGENT INFORMATION:
59 : NAME: NEWMARK, Sheridan
60 : REGISTRATION NUMBER: 20,620
61 : REFERENCE/KEY NUMBER: MARUTA 2
62 : TELECOMMUNICATION INFORMATION:
63 : TELEPHONE: 202 628 5197
64 : TELEFAX: 202 737 3528
65 : TELEX: 248633
66 : INFORMATION FOR SEQ ID NO: 1:
67 : SEQUENCE CHARACTERISTICS:
68 : LENGTH: 10 amino acids
69 : TYPE: amino acid
70 : STRANDEDNESS: single
71 : TOPOLOGY: linear
72 : MOLECULE TYPE: peptide
73 : US 09 341 829-1
74
75 Query Match
76 Best Local Similarity: 63.9%; Score 23; DB 1; Length 10;
77 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
78
79 QY 2 EAGR 6
80 :||||
81 DB 5 QGAGR 9
82
83 RESULT 14
84 US 09 341 829-1
85 : Sequence 1, Application US/0844352
86 : Patent No. 5591612
87 : GENERAL INFORMATION:
88 : APPLICANT: MARUTA, Kazuhiko
89 : APPLICANT: KUROTA, Michio
90 : APPLICANT: SUGIMOTO, Toshiyuki
91 : APPLICANT: MIYAKE, Toshio
92 : TITLE OF INVENTION: TRIHALOGENE-RELEASING ENZYME, AND ITS
93 : TITLE OF INVENTION: PREPARATION AND USES
94 : NUMBER OF SEQUENCES: 11
95 : CORRESPONDENCE ADDRESS:
96 : ADDRESSEE: BROWDY AND NEWMARK
97 : STREET: 419 Seventh Street, N.W., Suite 400
98 : CITY: Washington
99 : STATE: D.C.
100 : COUNTRY: USA
101 : ZIP: 20004
102 : COMPUTER READABLE FORM:
103 : MEDIUM TYPE: Floppy disk
104 : COMPUTER: IBM PC compatible
105 : OPERATING SYSTEM: PC DOS/MS DOS
106 : SOFTWARE: Patent In Release #1.0, Version #1.25
107 : CURRENT APPLICATION DATA:
108 : APPLICATION NUMBER: US/08/443,352
109 : FILING DATE: 17-MAY-1995
110 : CLASSIFICATION: 435
111 : PRIOR APPLICATION DATA:
112 : APPLICATION NUMBER: US 08/253,171
113 : FILING DATE: 02 JUN-1994
114 : APPLICATION NUMBER: JP 15638/1994
115 : FILING DATE: 24 JUN-1994
116 : PRIOR APPLICATION DATA:
117 : APPLICATION NUMBER: JP 43343/1994
118 : FILING DATE: 09 DEC-1993
119 : PRIOR APPLICATION DATA:
120 : APPLICATION NUMBER: JP 79291/1994
121 : FILING DATE: 28 MAR-1994
122 : ATTORNEY/AGENT INFORMATION:
123 : NAME: NEWMARK, Sheridan
124 : REGISTRATION NUMBER: 20,620
125 : REFERENCE/KEY NUMBER: MARUTA 2
126 : TELECOMMUNICATION INFORMATION:
127 : TELEPHONE: 202 628 5197
128 : TELEFAX: 202 737 3528
129 : TELEX: 248633
130 : INFORMATION FOR SEQ ID NO: 1:
131 : SEQUENCE CHARACTERISTICS:
132 : LENGTH: 10 amino acids
133 : TYPE: amino acid
134 : STRANDEDNESS: single
135 : TOPOLOGY: linear
136 : MOLECULE TYPE: peptide
137 : US 09 341 829-1

```

QY 1 REGAR 6
 1111
 1111

RESULT 14

5182464 10

PATENT NO. 5182464

APPLICANT: USER, CHERIL A. CURTISS, LINDA K. SMITH, RICHARD

TITLE OF INVENTION: POLYPEPTIDE ANALOGS OF APOLIPOPROTEIN E

NUMBER OF SEQUENCES: 14

CURRENT APPLICATION DATA:

APPLICANT: USER, CHERIL A. CURTISS, LINDA K. SMITH, RICHARD

APPLICANT: USER, CHERIL A. CURTISS, LINDA K. SMITH, RICHARD

APPLICANT: USER, CHERIL A. CURTISS, LINDA K. SMITH, RICHARD

APPLICANT: USER, CHERIL A. CURTISS, LINDA K. SMITH, RICHARD

APPLICANT: USER, CHERIL A. CURTISS, LINDA K. SMITH, RICHARD

APPLICANT: USER, CHERIL A. CURTISS, LINDA K. SMITH, RICHARD

APPLICANT: USER, CHERIL A. CURTISS, LINDA K. SMITH, RICHARD

APPLICANT: USER, CHERIL A. CURTISS, LINDA K. SMITH, RICHARD

APPLICANT: USER, CHERIL A. CURTISS, LINDA K. SMITH, RICHARD

APPLICANT: USER, CHERIL A. CURTISS, LINDA K. SMITH, RICHARD

APPLICANT: USER, CHERIL A. CURTISS, LINDA K. SMITH, RICHARD

APPLICANT: USER, CHERIL A. CURTISS, LINDA K. SMITH, RICHARD

APPLICANT: USER, CHERIL A. CURTISS, LINDA K. SMITH, RICHARD

APPLICANT: USER, CHERIL A. CURTISS, LINDA K. SMITH, RICHARD

APPLICANT: USER, CHERIL A. CURTISS, LINDA K. SMITH, RICHARD

APPLICANT: USER, CHERIL A. CURTISS, LINDA K. SMITH, RICHARD

APPLICANT: USER, CHERIL A. CURTISS, LINDA K. SMITH, RICHARD

APPLICANT: USER, CHERIL A. CURTISS, LINDA K. SMITH, RICHARD

APPLICANT: USER, CHERIL A. CURTISS, LINDA K. SMITH, RICHARD

APPLICANT: USER, CHERIL A. CURTISS, LINDA K. SMITH, RICHARD

APPLICANT: USER, CHERIL A. CURTISS, LINDA K. SMITH, RICHARD

APPLICANT: USER, CHERIL A. CURTISS, LINDA K. SMITH, RICHARD

APPLICANT: USER, CHERIL A. CURTISS, LINDA K. SMITH, RICHARD

APPLICANT: USER, CHERIL A. CURTISS, LINDA K. SMITH, RICHARD

APPLICANT: USER, CHERIL A. CURTISS, LINDA K. SMITH, RICHARD

APPLICANT: USER, CHERIL A. CURTISS, LINDA K. SMITH, RICHARD

APPLICANT: USER, CHERIL A. CURTISS, LINDA K. SMITH, RICHARD

APPLICANT: USER, CHERIL A. CURTISS, LINDA K. SMITH, RICHARD

APPLICANT: USER, CHERIL A. CURTISS, LINDA K. SMITH, RICHARD

APPLICANT: USER, CHERIL A. CURTISS, LINDA K. SMITH, RICHARD

APPLICANT: USER, CHERIL A. CURTISS, LINDA K. SMITH, RICHARD

APPLICANT: USER, CHERIL A. CURTISS, LINDA K. SMITH, RICHARD

APPLICANT: USER, CHERIL A. CURTISS, LINDA K. SMITH, RICHARD

APPLICANT: USER, CHERIL A. CURTISS, LINDA K. SMITH, RICHARD

APPLICANT: USER, CHERIL A. CURTISS, LINDA K. SMITH, RICHARD

APPLICANT: USER, CHERIL A. CURTISS, LINDA K. SMITH, RICHARD

APPLICANT: USER, CHERIL A. CURTISS, LINDA K. SMITH, RICHARD

APPLICANT: USER, CHERIL A. CURTISS, LINDA K. SMITH, RICHARD

APPLICANT: USER, CHERIL A. CURTISS, LINDA K. SMITH, RICHARD

APPLICANT: USER, CHERIL A. CURTISS, LINDA K. SMITH, RICHARD

APPLICANT: USER, CHERIL A. CURTISS, LINDA K. SMITH, RICHARD

APPLICANT: USER, CHERIL A. CURTISS, LINDA K. SMITH, RICHARD

APPLICANT: USER, CHERIL A. CURTISS, LINDA K. SMITH, RICHARD

APPLICANT: USER, CHERIL A. CURTISS, LINDA K. SMITH, RICHARD

APPLICANT: USER, CHERIL A. CURTISS, LINDA K. SMITH, RICHARD

APPLICANT: USER, CHERIL A. CURTISS, LINDA K. SMITH, RICHARD

APPLICANT: USER, CHERIL A. CURTISS, LINDA K. SMITH, RICHARD

APPLICANT: USER, CHERIL A. CURTISS, LINDA K. SMITH, RICHARD

APPLICANT: USER, CHERIL A. CURTISS, LINDA K. SMITH, RICHARD

APPLICANT: USER, CHERIL A. CURTISS, LINDA K. SMITH, RICHARD

APPLICANT: USER, CHERIL A. CURTISS, LINDA K. SMITH, RICHARD

APPLICANT: USER, CHERIL A. CURTISS, LINDA K. SMITH, RICHARD

APPLICANT: USER, CHERIL A. CURTISS, LINDA K. SMITH, RICHARD

APPLICANT: USER, CHERIL A. CURTISS, LINDA K. SMITH, RICHARD

APPLICANT: USER, CHERIL A. CURTISS, LINDA K. SMITH, RICHARD

APPLICANT: USER, CHERIL A. CURTISS, LINDA K. SMITH, RICHARD

APPLICANT: USER, CHERIL A. CURTISS, LINDA K. SMITH, RICHARD

APPLICANT: USER, CHERIL A. CURTISS, LINDA K. SMITH, RICHARD

APPLICANT: USER, CHERIL A. CURTISS, LINDA K. SMITH, RICHARD

APPLICANT: USER, CHERIL A. CURTISS, LINDA K. SMITH, RICHARD

APPLICANT: USER, CHERIL A. CURTISS, LINDA K. SMITH, RICHARD

STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-09-471-7800-85

Query Match: 63.9%; Score: 23; DB: 1; Length: 14;
 Best Local Similarity: 66.7%; Pred. No.: 73;
 Matches: 4; Conservative: 1; Mismatches: 0; Gaps: 0;

QY 1 REGAR 6
 1111
 1111

DB 3 REGAR 8

Search completed: May 1, 2001, 11:27:42
 Job time: 152 sec

Query Match: 63.9%; Score: 23; DB: 6; Length: 10;
 Best Local Similarity: 66.7%; Pred. No.: 52;
 Matches: 5; Conservative: 0; Mismatches: 1; Indels: 0; Gaps: 0;

QY 1 REGAR 6
 1111
 1111

RESULT 15

US-09-471-7800-85

Sequence 85, Application US/084717800

Patent No. 5182464

GENERAL INFORMATION:

APPLICANT: Casterman, Cecile

APPLICANT: Hamers, Raymond

TITLE OF INVENTION: Immunoglobulins Devoid of Light Chains

NUMBER OF SEQUENCES: 140

CORRESPONDENCE ADDRESS:

ADDRESS: Finnmark, Henderson, Farabow, Garrett & Dunnet

STREET: 1400 I Street, N.W.

CITY: Washington

STATE: D.C.

COUNTRY: USA

ZIP: 20005 3315

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: 09/084717800

FILING DATE: 06-JUN-1995

CLASSIFICATION: 530

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/084717800

FILING DATE: 17-AUG-1994

APPLICATION NUMBER: FR 92402426.0

FILING DATE: 21-AUG-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: FR 93401410.4

FILING DATE: 21-MAY-1993

ATTORNEY/AGENT INFORMATION:

NAME: Potter, Jane E.R.

REGISTRATION NUMBER: 34,382

REFERENCE/WORK NUMBER: 04958.0008-00000

TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-408-4000

TELEFAX: 202-408-4400

INFORMATION FOR SEQ ID NO: 85:

SEQUENCE CHARACTERISTICS:

LENGTH: 14 amino acids

TYPE: amino acid



Genware version 4.5
Copyright (c) 1993 - 2000 Genware Ltd.

us-09-341-829a-5_copy_142_148

us-09-341-829a-5_copy_142_148
18,646, 11,24,15 : Search time 25.6 seconds
(without alignments)

us-09-341-829a-5_copy_142_148

us-09-341-829a-5_copy_142_148

us-09-341-829a-5_copy_142_148

us-09-341-829a-5_copy_142_148

us-09-341-829a-5_copy_142_148

us-09-341-829a-5_copy_142_148

us-09-341-829a-5_copy_142_148

us-09-341-829a-5_copy_142_148

us-09-341-829a-5_copy_142_148

us-09-341-829a-5_copy_142_148

us-09-341-829a-5_copy_142_148

us-09-341-829a-5_copy_142_148

us-09-341-829a-5_copy_142_148

us-09-341-829a-5_copy_142_148

us-09-341-829a-5_copy_142_148

us-09-341-829a-5_copy_142_148

us-09-341-829a-5_copy_142_148

us-09-341-829a-5_copy_142_148

us-09-341-829a-5_copy_142_148

us-09-341-829a-5_copy_142_148

us-09-341-829a-5_copy_142_148

us-09-341-829a-5_copy_142_148

us-09-341-829a-5_copy_142_148

us-09-341-829a-5_copy_142_148

us-09-341-829a-5_copy_142_148

us-09-341-829a-5_copy_142_148

us-09-341-829a-5_copy_142_148

us-09-341-829a-5_copy_142_148

us-09-341-829a-5_copy_142_148

us-09-341-829a-5_copy_142_148

us-09-341-829a-5_copy_142_148

us-09-341-829a-5_copy_142_148

us-09-341-829a-5_copy_142_148

us-09-341-829a-5_copy_142_148

us-09-341-829a-5_copy_142_148

us-09-341-829a-5_copy_142_148

us-09-341-829a-5_copy_142_148

us-09-341-829a-5_copy_142_148

us-09-341-829a-5_copy_142_148

us-09-341-829a-5_copy_142_148

us-09-341-829a-5_copy_142_148

us-09-341-829a-5_copy_142_148

us-09-341-829a-5_copy_142_148

us-09-341-829a-5_copy_142_148

us-09-341-829a-5_copy_142_148

us-09-341-829a-5_copy_142_148

us-09-341-829a-5_copy_142_148

us-09-341-829a-5_copy_142_148

us-09-341-829a-5_copy_142_148

us-09-341-829a-5_copy_142_148

us-09-341-829a-5_copy_142_148

us-09-341-829a-5_copy_142_148

us-09-341-829a-5_copy_142_148

us-09-341-829a-5_copy_142_148

us-09-341-829a-5_copy_142_148

us-09-341-829a-5_copy_142_148

us-09-341-829a-5_copy_142_148

us-09-341-829a-5_copy_142_148

us-09-341-829a-5_copy_142_148.sizlim.rpr

us-09-341-829a-5_copy_142_148.sizlim.rpr

us-09-341-829a-5_copy_142_148.sizlim.rpr

us-09-341-829a-5_copy_142_148.sizlim.rpr

us-09-341-829a-5_copy_142_148.sizlim.rpr

```

RESULT 5
SI-614
ribosomal protein l16 - Bacillus stearothermophilus (fragment)
C:Species: Bacillus stearothermophilus
C:Date: 30-Jun-1991 #sequence_revision 40-Jun-1991 #text_change 18-Aug-1999
C:Accession: S10614
R:Kramer, W.J.; Hatakeyama, T.; Kimura, M.
Biol. Chem. Hoppe-Seyler 371, 641-636, 1990
A:Title: Nucleotide sequences of Bacillus stearothermophilus ribosomal protein genes: pa
A:Reference number: S10610; M010:91025633
A:Accession: S10614
A:Status: not computed with conceptual translation
A:Molecule type: RNA
A:Residues: 1-26, 800.
A:Suprafamily: Escherichia coli ribosomal protein L16
C:Keywords: protein biosynthesis; ribosome

Query Match 63.9% Score 23; DB 2; Length 26;
Best Local Similarity 71.4% Pred. No. 1, 5e-02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 REWARD 7
DB 11 KKHCHM 17

RESULT 4
S21276
glutathione transferase (P2.5.1.18) alpha-yr - rat (fragment)
C:Species: Rattus norvegicus (Norway rat)
C:Date: 22-Nov-1994 #sequence_revision 30-Jan-1998 #text_change 30-Jan-1998
C:Accession: S21276
R:Kikuchi, T.; Iuchiya, T.; Shikata, Y.; Sasaki, F.; Tadaya, O.; Horie, T.; Satoh, T.
Biochem. J. 283, 307-311, 1992
A:Title: Developmental aspects of a unique glutathione S-transferase subunit Ys in the
Yotus.
A:Reference number: S21276, M010:92231842
A:Accession: S21276
A:Molecule type: protein
A:Residues: 1-28, 414
C:Superfamily: glutathione transferase
C:Keywords: dimer; liver; transferase

Query Match 63.9% Score 20; DB 2; Length 26;
Best Local Similarity 66.7% Pred. No. 1, 5e-02;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 REWARD 7
DB 10 KKHCHM 15

RESULT 5
A49480
major immunophilin hsp56 - chicken (fragment)
C:Species: Gallus gallus (chicken)
C:Date: 06-Oct-1994 #sequence_revision 18-Nov-1994 #text_change 28-Apr-1995
C:Accession: A49480
R:Yam, A.W.; Boudreau, L.M.; Levine, J.W.; Heinriksson, R.L.; Heibel Jr., M.K.
Biochemistry 32, 12571-12576, 1993
A:Title: An avian p56c-biogenic domain of 17,000 daltons is isolated following limited
A:Reference number: A49480; M010:94072550
A:Accession: A49480
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-15, 470
A:Experimental source: chicken
A:Keywords: protein biosynthesis; ribosome

```

```

Query Match 61.1% Score 22; DB 2; Length 15;
Best Local Similarity 80.0% Pred. No. 1, 5e-02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 REWAG 5
DB 2 REWTC 6

RESULT 6
A44598
endo-1,4 beta xylanase (E3.1.2.1.6) Thielavia terrestris (strain Z75a) (fragment)
C:Species: Thielavia terrestris
C:Date: 04-Mar-1997 #sequence_revision 04-Mar-1997 #text_change 26-Feb-1998
C:Accession: A44598
R:Yaquich, M.
submitted to the Protein Sequence Database, March 1994
A:Reference number: A44593
A:Accession: A44598
A:Status: preliminary
A:Molecule type: Protein
A:Residues: 1-30, 574
C:Note: strain 255b
C:Superfamily: endo-1,4 beta xylanase, endo-1,4 beta xylanase homology
C:Keywords: glycosidase; hydrolase

```

```

Query Match 61.1% Score 22; DB 2; Length 30;
Best Local Similarity 80.0% Pred. No. 3e-02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 REWAG 6
DB 21 REWAG 25

RESULT 7
150633
c-rel protein - chicken (fragment)
C:Species: Gallus gallus (chicken)
C:Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 21-Jul-2000
C:Accession: 150633
R:Hannink, M.; Temin, H.M.
Oncogene 5, 1843-1850, 1990
A:Title: Structure and autoregulation of the c-rel promoter.
A:Reference number: 150633; M010:9114746
A:Accession: 150633
A:Status: preliminary; translated from cDNA/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-9, <NAN>
A:Cross references: EMBL: X56443, M010:963338, FIDB: CAA09822.1, FIDB: 584483

```

```

Query Match 58.3% Score 21; DB 2; Length 7;
Best Local Similarity 100.0% Pred. No. 2e-05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GAGR 6
DB 6 GAGR 9

RESULT 8
924745
glutathione transferase (P2.5.1.18) 1-2 - rat (fragment)
C:Species: Rattus norvegicus (Norway rat)
C:Date: 25-Oct-1987 #sequence_revision 25-Oct-1987 #text_change 18-Jun-1993
C:Accession: B24735
R:Mannervik, B.; Auld, P.; Gauthier, C.; Jonsson, B.; Tabler, M.K.; Warholm, M.; Jon
sson, M.; Asadi, S.; G. A. 82, 7292-7295, 1985
A:Title: Identification of three classes of cytosolic glutathione transferase common
A:Reference number: A24745; M010:86042634

```


A:Accession: B04347
 A:Molecule type: Protein
 A:Superfamily: 118 - NAR
 C:Keywords: glutathione transferase

Query Match 58.0%; Score 21; DB 2; Length 18;

Best Local Similarity 80.0%; Prod. No. 4.6e+02;

Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 GAGRM 7

DB 11 GIGRM 15

RESULT 7

PROTEIN 7

C:Species: Mus musculus (house mouse)

C:Date: 28 Feb 1999 #sequence_revision 28 Feb 1999 #text_change 30-May-1997

C:Accession: S04505

C:Author: A.; Mollnes, S.; Boud, L.

C:Nature 416, 842-846, 1985

A:Title: Genomic organization of the genes encoding mouse T-cell receptor alpha-chain.

A:Reference number: EMBL:85296342

A:Accession: S04505

A:Molecule type: DNA

A:Residues: 1-20 - WTG

A:Cross reference: EMBL:X02859

A:Note: This sequence was determined from the germ-line gene

C:Keywords: T cell receptor

Query Match 58.0%; Score 21; DB 2; Length 20;

Best Local Similarity 66.0%; Prod. No. 3.9e+02;

Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 GAGRM 7

DB 6 GAGRM 8

RESULT 10

PROTEIN 10

C:Species: Homo sapiens (man)

C:Date: 27 Oct 1997 #sequence_revision not Rev 1996 #text_change 21 Jan 2000

C:Accession: S04546

C:Author: F.H.; Sorenson, F.S.; Peterson, T.E.; Ellmann, J.; Rasmussen, L.R.

C:Nature 370, 91-94, 1995

A:Title: Polymorphisms in the synuclein consensus motif of the alpha-synuclein fragment, NAC

A:Reference number: S04546; M01:95474376

A:Accession: S04546

A:Molecule type: Preliminary

A:Molecule type: Protein

A:Residues: 1-28 - EN

Query Match 58.0%; Score 21; DB 2; Length 28;

Best Local Similarity 100.0%; Prod. No. 4.6e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 GAGRM 7

DB 24 GAGRM 28

RESULT 11

PROTEIN 11

C:Species: Pseudomonas syringae

C:Date: 05 Apr 1994 #sequence_revision 05 Apr 1994 #text_change 07-May 1996

C:Accession: PX0069
 C:Species: Bacillus subtilis (B. subtilis)
 C:Date: 114, 940-945, 1993

A:Title: An inducible NADP-dependent D-phenylserine dehydrogenase from Pseudomonas

A:Reference number: PX0069; M01:94186519

A:Accession: PX0069

A:Molecule type: Protein

A:Residues: 1-28 - PAK

A:Comment: This enzyme consists of two identical subunits. It catalyzes the oxidation

C:Keywords: NADP; oxidoreductase

C:Date: 13-18/Region: NADP binding status predicted

Query Match 58.0%; Score 21; DB 2; Length 28;

Best Local Similarity 80.0%; Prod. No. 4.6e+02;

Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 GAGRM 7

DB 13 GIGRM 17

RESULT 12

PROTEIN 12

C:Species: Fusarium sporotrichoides

C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 06-Jun-1997

C:Accession: PA0051

C:Author: L.P.; Fukaya, N.; Suduta, Y.; Oono, Y.; Ishihara, K.; Isidori, A.

C:Nature 370, 91-94, 1995

A:Reference number: PA0051

A:Accession: PA0051

A:Molecule type: Protein

A:Residues: 1-15 - QPQ

Query Match 55.0%; Score 20; DB 2; Length 18;

Best Local Similarity 80.0%; Prod. No. 4.1e+02;

Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 REGM 5

DB 8 REGM 12

RESULT 13

PROTEIN 13

C:Species: Xenopus laevis (African clawed frog)

C:Date: 21 Nov 1993 #sequence_revision not Rev 1996 #text_change 20-Jun-2000

C:Accession: S04541; C45021

C:Author: J.; Grossberger, D.; de Tassier, L.

C:Nature 370, 91-94, 1995

A:Title: Organization and rearrangement of immunoglobulin M genes in the amphibian X

A:Reference number: S04541; M01:95474376

A:Accession: S04541

A:Molecule type: DNA

A:Residues: 1-17 - SCH

A:Cross reference: EMBL:X14918; M01:944805; M01:944805; M01:944805

A:Author: P.N.; Amey, J.; Suzuki, D.; Littman, G.W.

C:Nature 370, 91-94, 1995

A:Title: Eleven distinct V-B gene families and additional patterns of sequence variation

A:Reference number: M01:944805; M01:944805

A:Accession: S04541

A>Status: preliminary; not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 1-17 - SHAL

C:Superfamily: immunoglobulin V region; immunoglobulin homology

C:Keywords: heterotetramer; immunoglobulin

Query Match 55.0%; Score 20; DB 2; Length 17;

Best Local Similarity 80.0%; Prod. No. 4.6e+02;

Matches 4: Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 GAGRM 7
II 1
IR 6 GAGTM 14

RESULT 14

A33361
NAME: regulated phosphoprotein, 21K - rat (fragment)
Species: Rattus norvegicus (Norway rat)
Release date: 1999 #sequence_revision 08 Dec 1989 #text_change 25-Oct-1996
Accession: A33361
Kleinman J.F., H.C.; Girault, J.A.; Williams, K.R.; LoPresti, M.B.; Greengard, P.
J. Biol. Chem. 264, 7726-7734, 1989
Affili: App-21, a cyclic AMP-regulated phosphoprotein (Mr 21,000) enriched in depamin-
etic studies of its phosphorylation in vitro.
Accession number: A33361; MIM:89214228
Accession: A33361
Status: preliminary
Molecular type: protein
Accession: 1-19 -AF
Keywords: phosphoprotein

Query Match 55.6%; Score 20; DB 2; Length 19;

Best Local Similarity 50.0%; Pred. No. 5,100,02;
Matches 4: Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 REGAGK 6
II 1
IR 11 KSGAGK 16

RESULT 15

S14161
Probable ribosomal protein, 5.1K protein
Accession: S14161
Species: Saccharopolyspora erythraea
Release date: 1997 #sequence_revision 10 Oct-1997 #text_change 21 Nov-1998
Accession: S14161
Kistritzky, P.; Green, R.; Packman, L.C.; Rawlings, R.J.; Staunton, J.; Leadlay, P.F.
Eur. J. Biochem. 195, 823-830, 1991
Affili: An acyl-carrier-protein - thioesterase domain from the 6-deoxyerythronolide B S-
cell.

Accession number: S14091; MIM:91154424

Accession: S14161

Molecular type: protein

Accession: 1-20 -CAF

Accession: 1-20 -CAF

Accession: 1-20 -CAF

Accession: 1-20 -CAF

Accession: 1-20 -CAF

Accession: 1-20 -CAF

Accession: 1-20 -CAF

Accession: 1-20 -CAF

Accession: 1-20 -CAF

Accession: 1-20 -CAF

Accession: 1-20 -CAF

Accession: 1-20 -CAF

Accession: 1-20 -CAF

Accession: 1-20 -CAF

Accession: 1-20 -CAF

Accession: 1-20 -CAF

Accession: 1-20 -CAF

Accession: 1-20 -CAF

Accession: 1-20 -CAF

Accession: 1-20 -CAF

Accession: 1-20 -CAF

Accession: 1-20 -CAF

Accession: 1-20 -CAF

Accession: 1-20 -CAF

Accession: 1-20 -CAF

Accession: 1-20 -CAF

Accession: 1-20 -CAF

Accession: 1-20 -CAF

Search completed: May 1, 2001, 11:26:20

Job time: 125 sec

GapUp version 4.5
Copyright (c) 1993-2000 CompuLink Ltd.

OR protein: protein source: distal sw model

Run on: May 1, 2001, 11:24:10 / Search time: 16.30 seconds
(without alignments)
14,977 Million cell updates/sec

LD file: US 09 341 829a 5 copy 142_148
protein source: 6
Sequence: 1 REFSEQEM 1

Standard table: B250862
GapUp 10.26 / GapExt 0.5

Score used: 9446 seqs, 3425446 residues

Total number of hits satisfying chosen parameters: 1767

Minimum hit seq length: 6

Maximum hit seq length: 40

Post processing: Maximum Match, 6%

Maximum Match, 10%

Listing first 45 summaries

LD base: SWISSprot 43*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	Description
1	24	63.9	20	1 RL16_BACSI	P23410 bacillus st
2	20	55.6	15	1 GDB3_34120	P81420 solanum tub
3	20	55.6	16	1 HBD2_C100A	P81443 clostridium
4	20	55.6	24	1 HPIA_PAB11	P14571 eryctolagus
5	20	55.6	26	1 E417_AL503	P11817 human adeno
6	20	55.6	28	1 EF05_SL10N	P80715 streptococ
7	14	52.8	14	1 SAE2_AB666	P11756 arabidopsi
8	19	52.8	15	1 F1RA_ANAF1	P12801 anas platyr
9	19	50.0	10	1 C9X0_FAT	P80432 rattus norv
10	19	50.0	10	1 C9X0_H1004	P80982 thomomys obe
11	19	50.0	24	1 KAD_BW111	P55140 bacillus th
12	19	50.0	25	1 NEH1_C110K	P44963 gallus galli
13	11	47.2	14	1 L414_F10ME	P01470 tremella mo
14	11	47.2	14	1 F1RA_B10SE	P14452 equus caball
15	11	47.2	15	1 F1RA_SYN7A	P14463 syncytus ca
16	11	47.2	16	1 F1RA_C10ST	P14535 ceratopter
17	11	47.2	16	1 F1RA_B100AS	P14449 equus asini
18	11	47.2	16	1 F1RA_F100A	P14556 felis c110e
19	11	47.2	16	1 F1RA_H110A	P14453 hylobates l
20	11	47.2	16	1 F1RA_M0110	P12803 macaca muse
21	11	47.2	16	1 F1RA_M101E	P14455 mandrillus
22	11	47.2	16	1 F1RA_G000E	P14459 odonocollus
23	11	47.2	16	1 F1RA_T101E	P14536 tapirus ter
24	11	47.2	17	1 F1RA_P101E	P14460 sus scrofa
25	11	47.2	18	1 F1RA_C100R	P14444 camelus dro
26	11	47.2	18	1 F1RA_L101L	P14454 lama glama
27	11	47.2	19	1 F1RA_B1010	P14441 boson bonas
28	11	47.2	19	1 F1RA_B0000	P14442 bubalus bub
29	11	47.2	19	1 F1RA_C100E	P14446 cervus elap
30	11	47.2	19	1 F1RA_C100E	P14447 cervus elap
31	11	47.2	19	1 F1RA_M100E	P14457 mont lucus m
32	11	47.2	19	1 F1RA_M100E	P14451 ovis aries
33	11	47.2	20	1 C00P_S100H	P80406 verticillu

ALIGNMENTS

RESULT 1

RL16_BACSI STANDARD: PKL: 25 AA.
AC P23410:
DT 01-NOV-1991 (Ref. 20, Created)
DT 01-NOV-1991 (Ref. 20, Last sequence update)
DT 01-NOV-1992 (Ref. 23, Last annotation update)
DE 50S RIBOSOMAL PROTEIN L16 (FRAGMENT).
GN RPLP.
OS Bacillus stearothermophilus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group.
OC Bacillus/Staphylococcus group; Bacillus.
OX NCBI_14810-1422;
RN [1]
RT SEQUENCE FROM N.A.
RX MEDLINE 91025643; PubMed 2222862;
FA Kikuchi W., Hatakeyama T., Kimura M.:
RT "Nucleotide sequences of Bacillus stearothermophilus ribosomal
protein genes: part of the ribosomal S10 operon."
RL Biol. Chem. Hoppe-Seyler 371:641-646(1990).
CC 1- FUNCTION: THIS PROTEIN BINDS BIRNBEY TO 23S RIBOSOMAL RNA AND IS
LOCATED AT THE A-SITE OF THE PEPTIDYLTRANSFERASE CENTER.
CC 1- SIMILARITY: BELONGS TO THE L16 FAMILY OF RIBOSOMAL PROTEINS.

CC THIS SWISS PROTEIN ENTRY IS COPYRIGHTED. IT IS PRODUCED THROUGH A COLLABORATION
BETWEEN THE SWISS INSTITUTE OF BIOINFORMATICS AND THE EMBL OUTSTATION
AT THE EUROPEAN BIOINFORMATICS INSTITUTE. THERE ARE NO RESTRICTIONS ON ITS
USE BY NON-PROFIT INSTITUTIONS AS LONG AS ITS CONTENT IS IN NO WAY
MODIFIED AND THIS STATEMENT IS NOT REMOVED. USAGE BY AND FOR COMMERCIAL
ENTITIES REQUIRES A LICENSE AGREEMENT. IF YOU ARE A USER OF THIS ENTRY, YOU
OR SEND AN EMAIL TO LICENSE@ISB.BE.

CC EMBL: X54994; CAA48741.1;
CC PIR: S10614; S10614.
CC InterPro: IPR000114;
CC Pfam: PF00252; Ribosomal_L16; 1.
CC PROSITE: PS00486; P10759RAL116; 1; LAF1AL.
CC PROSITE: PS00701; RIBOSOMAL_L16; 2; PAF1AL.
CC Ribosomal protein; tRNA binding
CC NON_TER 26 26
CC ALIGNMENT: 26 AA; 225 NW; SAF555367/101058 chr64;

Query Match 63.9%; Score 24; Len 1; Length 26;

Best local similarity 71.4%; Pval. No. 62;

Matches 5; Conservation 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 REGARM 7

Db 11 REGARM 17

RESULT 2

QY3_S0110 STANDARD: PKL: 15 AA.

Ac P81420:


```

EX  REFERENCE: 1209145; PubMed 6014721;
FA  Doolittle R.F., Schubert D., Schwartz S.A.;
EI  "Amino acid sequence studies on artiodactyl fibrinopeptides. I.
EI  Dromedary camel, mule deer, and cape buffalo.";
EL  Arch. Biochem. Biophys. 118:456-467(1967).
CC  -1- FUNCTION: FIBRINOGEN HAS A DOUBLE FUNCTION: YIELDING MONOMERS THAT
CC  POLYMERIZE INTO FIBRIN AND ACTING AS A COFACTOR IN PLATELET
CC  AGGREGATION.
CC  -1- SUBUNIT: HEXAMER (CONTAINING 2 SETS OF 3 IDENTICAL CHAINS
CC  (ALPHA, BETA, & GAMMA)), LINKED TO EACH OTHER BY DISULFIDE BONDS.
CC  -1- MISCELLANEOUS: CONVERSION OF FIBRINOGEN TO FIBRIN IS TRIGGERED BY
CC  THROMBIN, WHICH CLEAVES FIBRINOPEPTIDES A AND B FROM ALPHA & BETA
CC  CHAINS, AND THIS EXPOSES THE N-TERMINAL POLYMERIZATION SITES
CC  RESPONSIBLE FOR THE FORMATION OF THE SOFT CLOT.
KW  blood coagulation; Plasma.
EI  N_NTER 15 15
SL  SEQUENCE 15 AA. 1480 MW. 4E986A510B41076 CRC647

Query Match 47.28; Score 17; DB 1; Length 15;
Best Local Similarity 75.08; Pred. No. 8.5e+02;
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QT 2 E5W3 5
DB 11 A
DB 10 E5W3 14

```

Search completed: May 1, 2001, 11:25:50
 Job time: 100 sec

Query Match 37.88; Score 31; DB 18; Length 12;
 Best Local Similarity 100.00; Pred. No. 31;
 Matches 6; Conservation 0; Mismatches 0; Indels 0; Gaps 0;

QV 10 HKVSEQ 15
 111111
 16 4 HKVSEQ 9

RESULT 6
 Y60492
 Y60492 standard; Protein: 30 AA.
 XX
 KW Y60492;
 AC
 XX
 DT 01-JAN-1996 (first entry)
 XX
 DE Human normal bladder tissue EST encoded protein 164.
 XX
 KW human bladder treatment; gene expressed in human bladder cell line;
 KW cancer; gene therapy.
 XX
 OS Homo sapiens.
 XX
 PR 28-JUL-1999;
 XX
 PR 21-APR-1998; 98DE-1018620.
 XX
 PR 21-APR-1998; 98DE-1018620.
 XX
 PA (MELAN) METASTASIS OF SPERMATOCYTES MBL.
 XX
 PR Rosenthal A, Specht T, Hinzmann R, Schmitt A, Pillarsky C, Bahl R;
 XX
 PR WJL; 1999-002416/92.
 DR N-PSDB; 242208.
 XX
 XX New polypeptides and their nucleic acids, useful for treatment of
 PT Bladder tumour and identification of therapeutic agents
 XX
 PS Claim 24; Page 311; 866pp; German
 XX
 CC This invention describes novel polypeptide fragment sequences (I) and
 CC their encoding nucleic acids (II) which are highly expressed in normal
 CC bladder tissue and have cytostatic activity. (II) are used for
 CC treatment of bladder cancer and to isolate complete genes. (I) are
 CC used to identify agents suitable for the treatment of bladder tumours, to
 CC directly treat this form of cancer (including expression from gene
 CC therapy vectors), or are used in a preparation for cancer treatment. (I)
 CC is also used for the generation of specific antibodies. (II) are
 CC identified by assembling ESTs (expressed sequence tags) from a particular
 CC tissue type before comparison of expression patterns. This allows a
 CC significantly longer fragment of the gene to be revealed, and therefore
 CC reduces the number of failures because of ESTs from different libraries
 CC representing different parts of the same unknown gene distorting the
 CC estimated frequency of occurrence in a particular tissue. Y6029-Y6099
 CC represent protein fragments encoded by the human normal bladder tissue
 CC cDNA library derived EST fragments represented in 242122-242248.
 XX
 SL Sequence 30 AA;

Query Match 37.88; Score 31; DB 20; Length 30;
 Best Local Similarity 62.58; Pred. No. 82;
 Matches 6; Conservation 0; Mismatches 1; Indels 0; Gaps 0;

QV 6 RTHKRVSL 13
 111111
 16 24 RTHKRVSL 43

Query Match 36.64; Score 63; DB 16; Length 13;
 Best Local Similarity 84.38; Pred. No. 51;
 Matches 5; Conservation 0; Mismatches 1; Indels 0; Gaps 0;

QV 6 RTHKRVSL 11
 11111
 DB 7 RTHKRVSL 12

RESULT 8
 R67477
 R67477 standard; Peptide: 18 AA
 XX
 AC R67477;
 XX

Query Match 36.64; Score 63; DB 16; Length 13;
 Best Local Similarity 84.38; Pred. No. 51;
 Matches 5; Conservation 0; Mismatches 1; Indels 0; Gaps 0;

QV 6 RTHKRVSL 11
 11111
 DB 7 RTHKRVSL 12

RESULT 7
 R74800
 R74800 standard; Peptide: 17 AA.
 XX
 AC R74800;
 XX
 DT 09-JAN-1996 (first entry)
 XX
 DE Human tumour necrosis factor (hTNF) alpha (His 11).
 XX
 KW Interleukin-1 beta nonapeptide; IL-1 beta; fusion protein;
 KW human tumour necrosis factor alpha; hTNF alpha His 11; 163-171;
 KW irradiation protection; haematopoietic progenitor regeneration;
 KW immunostimulation; cancer cell growth reduction.
 XX
 OS Homo sapiens.
 XX
 PR Key Location/Qualifiers
 PR Peptide 1-13 "hTNF alpha His 11 residues 1-13
 PR (residues 14-157 are absent)"
 XX
 PR W09512675 A1.
 XX
 PR 11-MAY-1995.
 XX
 PR 03-NOV-1994; 94WO-0812656.
 XX
 PR 04-NOV-1993; 94US-0147673.
 XX
 PA (CYTO) CYTOTOXICAL PHARM INC.
 PA (CYTO) CYTOTOXICAL PHARMACEUTICALS INC.
 XX
 PR Bolton AP, Siddhu RS;
 XX
 PR WPI; 1995-185778/24.
 DR N-PSDB; Q90266.
 XX
 XX New fusion protein for use in tumour therapy comprising a tumour
 PT necrosis factor molecule and interleukin 1 beta residues 153-171.
 XX
 PS Claim 4; Fig 4; 44pp; English.
 XX
 CC Q90266 encodes R74800 hTNF alpha His 11 (early residues 1-13 given),
 CC which was fused to the IL-1 beta residues 163-171 nonapeptide to
 CC produce a fusion protein. The fusion protein can be used to enhance
 CC protection against irradiation, and to accelerate the regeneration
 CC of haematopoietic progenitors in animals treated with cytotoxic
 CC anti-cancer drugs and irradiation. In part, they can be used to
 CC immunostimulate patients that have undergone irradiation therapy,
 CC they may also be used to selectively reduce the growth of cancer
 CC cells.
 XX
 SL Sequence 13 AA;

D1 18 AUG 1995 (first entry)
 XX Mouse kappa/beta opoid receptor second intracellular loop fragment.
 XX
 XX Mouse kappa/beta mu opiod receptor; brain; primer; PCR; amplification; transmembrane domain; somatostatin receptor; human; expression vector; truncation; chimera; assay; probe.
 XX
 XX Synthesis.
 XX
 XX W0428132 A.
 XX
 XX 06 DEC 1994.
 XX
 XX 20 MAY 1994; 94WO 0505747.
 XX
 XX 20 MAY 1994; 94US 066246.
 XX
 XX 30 JUL 1994; 94US 010694.
 XX
 XX 05 NOV 1994; 94US 044792.
 XX
 XX (Abel) AB28 DEV 2496.
 XX
 XX 16 JUL 1994; Jelinek L, Yasuda K.
 XX
 XX WPI: 1994 02004204
 XX
 XX Peptide tubes and peptides derived from opiod receptor polypeptides for use in therapeutic compositions and in screening assays for useful drug substances.
 XX
 XX Based source: Euro Jec; 40pp; English.
 XX
 XX The amino acid sequence of the second intracellular loop of the mouse kappa/beta opiod receptors. The sequence has similarity to the second intracellular loop of the somatostatin receptor subunit 3 (67/676) the mouse kappa and delta receptors are 1 and 100. The second and third (94/674) loops of the opiod receptors are potentially involved in interaction with G proteins in a signal mechanism. Novel chimeric receptors were constructed where the second or third intracellular loops of the somatostatin receptor was recombined with the second and/or third loops of the kappa and/or delta mu opiod receptors. The role of these loops could be determined by observing whether the somatostatin receptors lost the ability to interact with G proteins or if the opiod receptors lost the ability to interact with G proteins. The chimeric receptors were constructed by recombination and ligation of PCR amplified fragments (see 975942-49 for primers). The chimeric opiod receptors thus produced are useful for the development of novel assays designed to select or improve substances, capable of interacting with the opiod receptor proteins, for use in diagnostics, drug design and therapeutic applications.
 XX
 XX Sequence: 18 AA;
 XX
 XX Query Match: 96.6%; Score 49; DB 16; Length 18;
 XX Best Local Similarity: 75.00%; Prod. No. 74;
 XX Matches: 6; Conservative: 0; Mismatches: 2; Indels: 0; Gaps: 0;
 XX
 XX 1 KAROLINRIKURSE 18
 XX 1 1 1 1
 XX 11 Kallitrip 18
 XX
 XX RESULT 9
 XX Y99947 standard; Protein: 20 AA.
 XX
 XX Y99947
 XX
 XX 28 JAN 2000 (first entry)
 XX
 XX Human myometrium tumour EST encoded protein 17.
 XX

KW Myometrium; tumour; human; expressed sequence tag; EST; uterine myoma treatment; carcinoma; cancer; gene therapy.
 XX
 XX Homo sapiens.
 XX
 XX DE19817947-A1.
 XX
 XX 28 OCT-1999.
 XX
 XX 17 APR-1998; 98DE 1017947.
 XX
 XX 17 APR 1998; 98DE 1017947.
 XX
 XX (META) METAFER GTS GENE-DEVELOPMENT MBL.
 XX
 XX Rosenthal A, Specht T, Hinzmann K, Schmitt A, Fritsky P, Dahl M.
 XX
 XX WPI: 1999-602480/52.
 XX
 XX N-PSDB: Z41980.
 XX
 XX New nucleic acid sequences expressed in uterine myoma, and derived polypeptides, for treatment of uterine carcinoma and identification of therapeutic agents
 XX
 XX Claim 23; Page 71; 86pp; German.
 XX
 XX This invention describes novel polypeptide sequences (I). Fragments of (I) fragments and their encoding nucleic acids (II) which are highly expressed in human uterine myoma. (II) are used for recombinant expression of (I) and to isolate complementary DNAs. (I) are used to identify agents suitable for treatment of uterine carcinoma, to directly treat this form of cancer (including expression from gene therapy vectors) and are used in a preparation for cancer treatment. (I) is also used for the generation of specific antibodies. (II) are identified by assembling ESTs (expressed sequence tags) from a particular tissue type before comparison of expression patterns. This allows a significantly longer fragment of the gene to be revealed and thereby reduces the number of failures associated with the fact that ESTs from different libraries may represent different parts of the same unknown gene. Also, disturbing the estimated frequency of occurrence in a particular tissue. Y9921 Y9949 represent protein fragments encoded by the human myometrium tumour cDNA library derived EST fragments represented in Z41950 Z41980.
 XX
 XX Sequence: 20 AA;
 XX
 XX Query Match: 96.6%; Score 30; DB 20; Length 20;
 XX Best Local Similarity: 50.00%; Prod. No. 80;
 XX Matches: 7; Conservative: 2; Mismatches: 0; Indels: 0; Gaps: 0;
 XX
 XX 1 KAROLINRIKURSE 14
 XX 1 1 1 1 1 1
 XX 6 Kwrtsrpkaklge 19
 XX
 XX RESULT 10
 XX P61303
 XX 10 P61303 standard; Protein: 22 AA
 XX
 XX P61303
 XX
 XX 19 JUN-1991 (first entry)
 XX
 XX Peptide acting selectively on the electric potential dependent sodium channel.
 XX
 XX Na channel; ATPase.
 XX
 XX Conus geographus.
 XX
 XX Key location/Qualifiers
 XX Modified-site: 6

```

F1 Modified-site 7 /label Hydroxy proline
F1 /label Hydroxy proline
F1 Modified site 17 /label Hydroxy proline
F1 /label Hydroxy proline
XX JP60239671-A.
XX
XX 28-NOV-1985.
XX
XX 15-MAY-1984; 84JP-0097070.
XX
XX 15-MAY-1984; 84JP-0097070.
XX
XX (MILUO) MITSUBISHI CHEM IND KK.
XX
XX WPI: 1986-017118/03.
XX
XX Reagent for analysing sodium channel in muscle cell of animal -
XX contains specified peptide acting selectively on the electric
XX potential dependent sodium channel.
XX
XX Claim 1: Page 417; 4pp; Japanese.
XX
XX Peptide may be incorporated into a reagent for the study of the
XX biochemical and pharmacological activities of Na channels in animal
XX muscle.
XX
XX Sequence 22 AA:

Query Match 96.5%; Score 40; DP 7; Length 22.
Best Local Similarity 42.9%; Pred. No. 89;
Matches 6; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 3 RRLKTHPKVSKQR 16
DB 1 11 1 1 1 1
1 rdeetpprkekdr 14

RESULT 11
P61304
11 P61304 standard; Protein: 22 AA.
XX
XX AC P61304;
XX
XX 19-JUN-1991 (first entry)
XX
XX Peptide acting selectively on the electric potential dependent sodium
XX channel.
XX
XX Na channel; ATPase.
XX
XX Cons geographicus.
XX
XX Key Location/Qualifiers
XX Modified-site 6 /label Hydroxy proline
XX Modified site 7 /label Hydroxy proline
XX Modified site 17 /label Hydroxy proline
XX Modified site 17 /label Hydroxy proline
XX
XX JP60239671-A.
XX
XX 28-NOV-1985.
XX
XX 15-MAY-1984; 84JP-0097070.
XX
XX 15-MAY-1984; 84JP-0097070.
XX
XX (MILUO) MITSUBISHI CHEM IND KK.
XX
XX WPI: 1986-017118/03.
XX
XX Reagent for analysing sodium channel in muscle cell of animal -
XX contains specified peptide acting selectively on the electric
XX potential dependent sodium channel.
XX
XX Claim 1: Page 417; 4pp; Japanese.
XX
XX Peptide may be incorporated into a reagent for the study of the
XX biochemical and pharmacological activities of Na channels in animal
XX muscle.
XX
XX Sequence 22 AA:

Query Match 96.5%; Score 40; DP 7; Length 22.

```

```

XX
XX WPI: 1986-017118/03.
XX
XX Reagent for analysing sodium channel in muscle cell of animal -
XX contains specified peptide acting selectively on the electric
XX potential dependent sodium channel.
XX
XX Claim 1: Page 417; 4pp; Japanese.
XX
XX Peptide may be incorporated into a reagent for the study of the
XX biochemical and pharmacological activities of Na channels in animal
XX muscle.
XX
XX Sequence 22 AA:

Query Match 96.5%; Score 40; DP 7; Length 22.
Best Local Similarity 42.9%; Pred. No. 89;
Matches 6; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 3 RRLKTHPKVSKQR 16
DB 1 11 1 1 1 1
1 rdeetpprkekdr 14

RESULT 12
P61305
10 P61305 standard; Protein: 22 AA.
XX
XX AC P61305;
XX
XX 19-JUN-1991 (first entry)
XX
XX Peptide acting selectively on the electric potential dependent sodium
XX channel.
XX
XX Na channel; ATPase.
XX
XX Cons geographicus.
XX
XX Key Location/Qualifiers
XX Modified-site 6 /label Hydroxy proline
XX Modified site 7 /label Hydroxy proline
XX Modified site 17 /label Hydroxy proline
XX Modified site 17 /label Hydroxy proline
XX
XX JP60239671-A.
XX
XX 28-NOV-1985.
XX
XX 15-MAY-1984; 84JP-0097070.
XX
XX 15-MAY-1984; 84JP-0097070.
XX
XX (MILUO) MITSUBISHI CHEM IND KK.
XX
XX WPI: 1986-017118/03.
XX
XX Reagent for analysing sodium channel in muscle cell of animal -
XX contains specified peptide acting selectively on the electric
XX potential dependent sodium channel.
XX
XX Claim 1: Page 417; 4pp; Japanese.
XX
XX Peptide may be incorporated into a reagent for the study of the
XX biochemical and pharmacological activities of Na channels in animal
XX muscle.
XX
XX Sequence 22 AA:

Query Match 96.5%; Score 40; DP 7; Length 22.

```


FF 12-AUG-1999: 990S-0148441.
 FF 13-AUG-1999: 990S-0148565.
 FF 14-AUG-1999: 990S-0148684.
 FF 15-AUG-1999: 990S-0149368.
 FF 16-AUG-1999: 990S-0149175.
 FF 17-AUG-1999: 990S-0149426.
 FF 18-AUG-1999: 990S-0149722.
 FF 19-AUG-1999: 990S-0149723.
 FF 20-AUG-1999: 990S-0149729.
 FF 21-AUG-1999: 990S-0149902.
 FF 22-AUG-1999: 990S-0149930.
 FF 23-AUG-1999: 990S-0150564.
 FF 24-AUG-1999: 990S-0150884.
 FF 25-AUG-1999: 990S-0151065.
 FF 26-AUG-1999: 990S-0151066.
 FF 27-AUG-1999: 990S-0151080.
 FF 28-AUG-1999: 990S-0151303.
 FF 29-AUG-1999: 990S-0151338.
 FF 30-AUG-1999: 990S-0151930.
 FF 01-SEP-1999: 990S-0152364.
 FF 02-SEP-1999: 990S-0153070.
 FF 03-SEP-1999: 990S-0153758.
 FF 04-SEP-1999: 990S-0154018.
 FF 05-SEP-1999: 990S-0154039.
 FF 06-SEP-1999: 990S-0154779.
 FF 07-SEP-1999: 990S-0155139.
 FF 08-SEP-1999: 990S-0155484.
 FF 09-SEP-1999: 990S-0155659.
 FF 10-SEP-1999: 990S-0156458.
 FF 11-SEP-1999: 990S-0156596.
 FF 12-SEP-1999: 990S-0157117.
 FF 13-SEP-1999: 990S-0157753.
 FF 14-SEP-1999: 990S-0157865.
 FF 15-SEP-1999: 990S-0158029.
 FF 16-SEP-1999: 990S-0158232.
 FF 17-SEP-1999: 990S-0158369.
 FF 18-SEP-1999: 990S-0158292.
 FF 19-SEP-1999: 990S-0159294.
 FF 20-SEP-1999: 990S-0159295.
 FF 21-SEP-1999: 990S-0159329.
 FF 22-SEP-1999: 990S-0159330.
 FF 23-SEP-1999: 990S-0159331.
 FF 24-SEP-1999: 990S-0159637.
 FF 25-SEP-1999: 990S-0159638.
 FF 26-SEP-1999: 990S-0159584.
 FF 27-SEP-1999: 990S-0160741.
 FF 28-SEP-1999: 990S-0160747.
 FF 29-SEP-1999: 990S-0160768.
 FF 30-SEP-1999: 990S-0160770.
 FF 01-OCT-1999: 990S-0160814.
 FF 02-OCT-1999: 990S-0160815.
 FF 03-OCT-1999: 990S-0160980.
 FF 04-OCT-1999: 990S-0160981.
 FF 05-OCT-1999: 990S-0160989.
 FF 06-OCT-1999: 990S-0161404.
 FF 07-OCT-1999: 990S-0161405.
 FF 08-OCT-1999: 990S-0161406.
 FF 09-OCT-1999: 990S-0161459.
 FF 10-OCT-1999: 990S-0161360.
 FF 11-OCT-1999: 990S-0161461.
 FF 12-OCT-1999: 990S-0161920.
 FF 13-OCT-1999: 990S-0161921.
 FF 14-OCT-1999: 990S-0161993.
 FF 15-OCT-1999: 990S-0162142.

Query Match: 35.48; Score 29; UB 20; Length 9;

Best Local Similarity 50.0%; Prod. No. 1.2e+02;

Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 2 ARDLTKHKVS 13

13 1 1 1 1 1 1

14 7 additional symbols 18

RESULT 14

Y55614

ID Y55614 standard; peptide; 9 AA.

XX Y55614;

AC Y55614;

XX 17-JAN-2000 (first entry)

XX HLA binding plu-1 peptide.

DE Human cancer-associated polypeptide plu-1; ovarian cancer vaccine;

XX breast cancer; immune response; cytotoxic T lymphocyte; immunogenic;

XX therapeutic; plu-1 antigen; breast tumor; tumor antigen; epitope.

XX Synthetic.

OS Homo sapiens.

XX W09949034-A1.

XX 40 SEP-1999.

XX 19-MAR-1999; 990S-0800866.

XX 20-MAR-1999; 990S-0805877.

XX (389) SPECIAL: CARP: RES: HEP: HLA: Y.

XX Taylor papadimitriou J;

XX WPI: 1999-591090/50.

XX New nucleic acid encoding the cancer associated polypeptide plu-1, for

XX diagnosis, treatment and prevention of cancer, especially of breast and

XX ovary.

XX Example 2; Fig 12; 17pp; English.

XX The invention relates to a human cancer-associated polypeptide plu-1. The

XX plu-1 polypeptide can be recombinantly expressed by standard recombinant

XX methodology. Detection of the plu-1 nucleic acid or the polypeptide is

XX used for the following: (i) diagnosis (including imaging) and prognosis

XX of, and determination of susceptibility to, cancer, specifically ovarian

XX or breast cancer; and (ii) treating cancer (by inducing an immune response

XX against cancer cells, e.g. as a vaccine, or by antisense inhibition).

XX Antisense derived from the polypeptide are used to generate activated

XX cytotoxic T lymphocytes or dendritic cells, for subsequent return to

XX the patient for treatment of cancer. The polypeptide may also be used to

XX identify inhibitors of plu-1 activity. Fragments of the polypeptide, and

XX antibodies raised against plu-1, are useful as assay and imaging agents,

XX also therapeutically (to induce an anti-idiotypic response or where

XX conjugated to cytotoxic agents). The plu-1 antigen is expressed more

XX commonly in breast tumors than some known tumor antigens. Sequences

XX Y55614 625 represent predicted peptides from the plu-1 polypeptide which

XX may bind to the human class I alleles B27, A2, A3 and A11.

XX Sequence 9 AA;

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX


```

XX 10 MAY 1996 (first entry)
XX 1. albicans enolase peptide #6.
XX
XX Candida albicans: enolase; lambda-qt11; serum; antibody; candidiasis;
XX Thrush.
XX
XX Synthesis.
XX
XX WO020962.A1.
XX
XX 05 OCT 1996.
XX
XX 27 MAY 1996: 95W0 AD09176.
XX
XX 25 MAY 1996: 94AD 0004742.
XX
XX (OY001) UNIV CORNELL TECHNOLOGY.
XX
XX Franklyn KM. Warrington 18;
XX WP1: 1996 01294/46.
XX
XX Protein or peptide reactive with Candida albicans induced antibodies
XX also DNA and vectors, used to treat, prevent or diagnose C.
XX albicans infections
XX
XX claim 1: page 26; 3/99; English.
XX
XX peptides 981710-25 are derived from the Candida albicans enolase protein
XX (p91709). the peptides can be used to diagnose, in a serum sample,
XX antibodies induced by a Candida albicans infection and hence the
XX infection itself. the gene encoding the enolase was isolated from a
XX Candida albicans cDNA expression library in lambda-qt11 using human sera from a
XX patient with a clinical history of vaginitis. The resultant clone
XX contained a 470 bp open reading frame encoding residues 94-249 (K81709).
XX This clone was used to screen a genome library and obtain the full
XX length clone.
XX
XX Sequence: 15 AA:
XX
Query Match: 35.4%; Score 29; 108 bp; Length 15;
Ident. Locn. Similarity: 62.5%; Ident. No. 892
Matches: 1; Substitutions: 2; Mismatches: 1; Indels: 0; Gaps: 0;

```

```

XX 2 ADDITIONAL
XX 1 111111
XX 4 00101011

```

Search completed: May 1, 2001, 11:25:33
 Job Time: 100 sec



GenCorp version 4.5
Copyright (c) 1993-2000 Compugen Ltd.

EM protein protein search using sw model

Run on: May 1, 2001, 11:26:21 ; Search time 25.8 seconds
(without alignments)
42.619 million cell updates/sec

Database: us-09-341-829a-5_copy_164_179
Sequence: 1 KAGAGGATPPKQVETPK 16

Scoring matrix: BLOSUM62
Gapop 10.0 ; Gapext 0.5

Searches: 1 word seqs, 6972245 residues

Total number of hits satisfying chosen parameters: 6159

Minimum hit seq length: 6
Maximum hit seq length: 40

Post processing: Minimum Match 90%
Listing first 40 summaries

Database: PIR 671*
1: P0111*
2: P0121*
3: P0131*
4: P0141*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No	Score	Query Match	Length	DB	Id	Description
1	40	96.6	22	1	MXKNZ	25.66kDa toxin GIIIB
2	39	96.4	25	2	JP0044	ribosomal protein
3	29	95.4	42	1	MXKNZ	no conotoxin GIIIB
4	28	94.1	24	2	154429	gene NE1 protein
5	27	92.9	40	2	SG6448	trimethylamine dehydrogenase
6	26	91.7	21	2	PN0382	sperm chromatin protein
7	25	90.5	22	1	MXKNZ	no conotoxin GIIIA
8	25	90.3	22	2	P0084	ribosomal protein
9	25	90.1	23	2	P0084	ribosomal protein
10	25	90.0	30	2	S07217	ribosomal protein
11	24	89.3	9	2	P07074	translation elongation factor
12	24	89.3	18	2	A56871	retinol-binding protein
13	24	89.3	22	2	552357	hypothetical protein
14	24	89.3	23	2	A04243	amyloid protein AL
15	24	89.3	28	2	A44244	hexokinase (B) 2.7
16	24	89.3	29	2	A42242	brain type creatin
17	24	89.3	28	2	A01243	retinol-binding protein
18	24	89.3	29	2	A54244	cytochrome c oxidase
19	24	89.3	29	2	559609	peptide lysozyme
20	24	89.0	15	2	A47146	topoisomerase I
21	24	89.0	29	2	150757	pulK protein (limp)
22	24	89.0	21	2	A27719	trypsin (EC 3.4.21)
23	24	89.0	23	2	A44524	pregnancy-specific protein
24	24	89.0	25	2	JP0064	ribosomal protein
25	24	89.0	25	2	S44291	HLA DRB1 exon2 gene
26	24	89.0	26	2	S26166	linP protein - Esc
27	24	89.0	27	2	A47295	homodomain protein
28	24	89.0	28	2	563552	95K protein - Euba
29	24	89.0	29	2	A61165	endometrial protease

30	24	26.8	40	2	S57227	ribosomal protein
31	22	26.8	14	2	S09244	phosphorylation protein
32	22	26.8	14	2	A60770	cell surface alpha
33	22	26.8	16	2	P00491	retinol-binding protein
34	22	26.8	20	2	JP0070	ribosomal protein
35	22	26.8	20	2	S18582	hypothetical protein
36	22	26.8	22	2	JP0071	ribosomal protein
37	22	26.8	22	2	S29423	cellular protein
38	22	26.8	22	2	P00337	cellular protein
39	22	26.8	24	2	S10909	deoxyribonuclease
40	22	26.8	24	2	S38766	peptide chain
41	22	26.8	26	2	A28108	atrial gland protein
42	22	26.8	28	1	G006SV	gene 9 protein
43	22	26.8	28	2	137539	hypothetical protein
44	22	26.8	28	2	JP01045	hypothetical protein
45	22	26.8	29	2	A34114	peptidylase (B) 3

ALIGNMENTS

RESULT 1

MXKNZ
no conotoxin GIIIB [validated] - cone shell (Conus neorapheus)
R. Altermann names: conotoxin II (GIFX II), myoconotoxin II
C. Specter; Conus neorapheus (neorapheo) cone
Genbank: 14 May 1993 #sequence_revision 14 Nov 1994 #text change 15 Sep 2000
Accession: A01787; B23579
RefSeq: S. Nakamura, H. Ohizumi, Y. Kobayashi, J. Hirata, Y.
FEBS Lett. 155, 277-280, 1983
A. Title: The amino acid sequences of homodimeric hydrophobic conotoxins I
A. Reference number: A91409; MIM:8321070
Accession: A01787
A. Molecule type: protein
A. Residues: 1-22 (SAT)
P. J. Biol. Chem. 260, 9280-9288, 1985
A. Title: Conus neorapheus toxins that discriminate between neuronal and muscle sodium
A. Reference number: A23579; MIM:8526106
Accession: B23579
A. Molecule type: protein
A. Residues: 1-22 (SAT)
K. Hill, J. M. Altemann, P. F. Crank, D. J.
Submitted to the Brookhaven Protein Data Bank, April 1996
A. Reference number: A65705; PDB: 6HIB
A. Content: annotation: conformation by (D)B NMR, residues 1-22
Biochemistry 35, 8824-8835, 1996
A. Title: Three dimensional solution structure of no conotoxin GIIIB, a specific block
A. Reference number: A58596; MIM:8628040
Accession: A58596
A. Content: annotation: conformation by (D)B NMR
C. Superfamily: no conotoxin
Keywords: mutated; carboxyl end; hydrophobic; myoconotoxin; sodium channel inhibition
P. 15, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100
P. 1/26 Modified site: 4 hydrophobic (no) status: experimental
P. 2/2 Modified site: mutated carboxyl end (Aia) status: experimental

Query Match: 36.0%, Score 40; DB 1; Length 22;
Best local similarity: 42.9%; Prod. No. 98;
Matches: 6; Conservative: 2; Mismatches: 6; Indels: 0; Gaps: 0;

Oy: 4 KOLTPPKVSEQR 16

Db: 1 KACCTPPPKVDRR 14

RESULT 2

JP0044
ribosomal protein L40 - Escherichia vulneris (H. Altemann)

C. Specter; Escherichia vulneris
C. Date: 10 Mar 1994 #sequence_revision 28 Oct 1994 #text change 03 May 1996

[illegible]

Sequence	Length	Score	Id	Length	Gaps
Query Match	40	56	26	40	0
Best Local Similarity	9	48		9	0
Matches	4	Conservative	4	Mismatches	0
1	KARLWPTPRK	11			
2	QVPTPTPT	1			
3	PPPTPTPT	1			
4	PPPTPTPT	2			
5					
6					
7					
8					
9					
10					
11					
12					
13					
14					
15					
16					
17					
18					
19					
20					
21					
22					
23					
24					
25					
26					
27					
28					
29					
30					
31					
32					
33					
34					
35					
36					
37					
38					
39					
40					
41					
42					
43					
44					
45					
46					
47					
48					
49					
50					
51					
52					
53					
54					
55					
56					
57					
58					
59					
60					
61					
62					
63					
64					
65					
66					
67					
68					
69					
70					
71					
72					
73					
74					
75					
76					
77					
78					
79					
80					
81					
82					
83					
84					
85					
86					
87					
88					
89					
90					
91					
92					
93					
94					
95					
96					
97					

ribosomal protein L18 - Halobacterium volcanum (Fragment)
 A:Alternate names: ribosomal protein L18
 A:Species: Halobacterium volcanum
 A:Date: 31-Mar-1992 sequence_revision 41-Mar-1992 text_change 18-Jun-1993
 A:Accession: A3084
 A:McDonald, J.
 submitted to the Protein Sequence Database, June 1990
 A:Reference number: A3084
 A:Accession: A3084
 A:Molecule type: protein
 A:Residues: 1-23 -MDP
 A:Experimental source: strain DSM 6752
 A:Superfamily: rat ribosomal protein L5
 A:Keywords: protein biosynthesis; ribosome

Query Match
 best local similarity 46.4% Prod. No. 7-0002
 Matches 4 Conservative 4 Mismatches 0 Gaps 0

QY 1 KARLQETHEK 11
 : 1 1 1 1 1
 IQ 12 KEFVEYDQ 22

RESUL 10
 S07217

A:ribosomal protein L18 [validated] Halobacterium salinarum (fragment)
 A:Alternate names: ribosomal protein H118
 A:Species: Halobacterium salinarum
 A:Date: 31-Mar-1992 sequence_revision 31-Mar-1992 text_change 21-Jul-2000
 A:Accession: S07217; B3084; B3084
 A:Smith, R., Matthews, A., Yarbello, M., Willett, J., Drenth, J., R.
 Fear, J., Bloembergen, R., 501-509, 1978
 A:Title: The 5 S RNA-ribosome complex from an extreme halophile, Halobacterium salinarum
 A:Reference number: S07217; M0119; 79645,279
 A:Accession: S07217
 A:Molecule type: protein
 A:Residues: 1-40 -SMI
 A:Note: The source is designated as Halobacterium salinarum
 A:Note: The protein is designated as ribosomal protein H118
 A:McDonald, J.
 submitted to the Protein Sequence Database, June 1990
 A:Reference number: A3084
 A:Accession: B3084
 A:Molecule type: protein
 A:Residues: 1-26 -MDP
 A:Experimental source: strain DSM 670
 A:Note: The source is designated as Halobacterium halobium
 A:Accession: B3084
 A:Molecule type: protein
 A:Residues: 1-23 -MDP
 A:Experimental source: strain DSM 6754
 A:Note: The source is designated as Halobacterium salinarum
 A:Note: a 66 kDa protein complex consisting of protein L5, protein L15
 A:Superfamily: rat ribosomal protein L5
 A:Keywords: protein biosynthesis; ribosome

```

Query Match 40.5% Score 25; Lb 2; Length 40;
Best Local Similarity 46.4%; Pred. No. 9.4e+02;
Matches 4; Conservative 4; Mismatch 4; Indel 13; Gaps 0;

QY 1 KADLRTKHK 11
      1 1 1 1 1 1;
Db 12 KAREVTHYHG 22

RESULT 11
pc/074
Translation of onatation factor cef-1 beta chain (mouse)
C-Species: Mus musculus (house mouse)
Date: 18 Aug 2000 #sequence revision 18 Aug 2000 #text change 18 Aug 2000

```

C:Accession: P07074
 K:Ushio, A.; Kawakami, T.; Uchida, T.; Sakai, T.; Kamo, M.; Matsui, T.; Watanabe, Y.
 Electrophoresis 21: 1853-1871, 2000
 A:Title: Proteomic analysis of mouse brain: two dimensional electrophoresis profiles of
 A:Reference number: P07072
 A:Accession: P07074
 A:Molecule type: protein
 A:Residues: 1-9 (18)
 A:Experimental source: strain 667g/96; Str, spleen, brain, cortex
 C:Keywords: brain; cerebral cortex

 Query Match: 29.4%; Score 24; DB 2; Length 9;
 Best Local Similarity 80.0%; Pred. No. 1.4e+02;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 Q7 4 QKRPB
 I0 4 QKRPB

 RESULT 12
 A:Accession: A56871
 C:Species: Sus scrofa domestica (domestic pig)
 C:Title: 05-Jan-1996 #sequence_revision 21-Jul-1996 #text_change 09-May-1997
 C:Accession: A56871
 K:Hallings-Mann, M.L.; Trout, W.F.; Roberts, R.M.
 J. Biol. Reprod. 48: 998-1005, 1993
 A:Title: Porcine uterine retinol-binding proteins are identical gene products to the sea
 A:Reference number: A56871; M010-9324392
 A:Accession: A56871
 A>Status: preliminary; not compared with conceptual translation
 A:Molecule type: mRNA
 A:Residues: 1-18 (18)
 A:Experimental source: endometrium
 A:Title: Sequence extracted from NBR1 backbone (NBR1E-132147)
 C:Superfamily: lipocalin; lipocalin homology

 Query Match: 29.4%; Score 24; DB 2; Length 18;
 Best Local Similarity 41.7%; Pred. No. 8.4e+02;
 Matches 5; Conservative 1; Mismatches 6; Indels 0; Gaps 0;
 Q7 4 QKRPB
 I0 2 QKRPB

 RESULT 13
 A:Accession: S52457
 C:Species: Homo sapiens (man)
 C:Title: 08-May-1995 #sequence_revision 21-Jul-1996 #text_change 07-May-1999
 C:Accession: S52457
 K:Artyukhina, A.; Leversha, H.A.; Ferguson, Smith, M.; Moschonas, M.K.
 Submitted to the EMBL data library, March 1994
 A:Description: A cosmid clone mapped to human chromosome 11p15 detects a highly restricted
 A:Reference number: S52455
 A:Accession: S52457
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-22 (22)
 A:Cross-references: EMBL:X72881; NID:q667002; PUD:CAA51393.1; PDB:q667004

 Query Match: 29.4%; Score 24; DB 2; Length 22;
 Best Local Similarity 57.1%; Pred. No. 1e+03;
 Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 Q7 10 HKVSPQR
 I0 10 HKVSPQR

RESULT 14
 A:Accession: A29323
 C:Species: Homo sapiens (man)
 C:Title: 21-Feb-1990 #sequence_revision 21-Feb-1990 #text_change 16-Aug-1996
 C:Accession: A29323
 K:Garcia, P.J.; Trull, P.C.; Wright, J.; Pras, M.; Prangione, R.
 J. Clin. Invest. 83: 836-843, 1989
 A:Title: Systemic senile amyloidosis. Identification of a new protein in (transferrin
 A:Reference number: A29323; M010-89155905
 A:Accession: A29324
 A>Status: preliminary
 A:Molecule type: protein
 A:Residues: 1-27 (27)
 C:Superfamily: immunoglobulin V region; immunoglobulin homology
 C:Keywords: heterotetramer; immunoglobulin

 Query Match: 29.4%; Score 24; DB 2; Length 27;
 Best Local Similarity 62.5%; Pred. No. 1.3e+03;
 Matches 5; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 Q7 7 TPKKVSSE 14
 I0 5 TPKKVSSE 12

 RESULT 15
 A:Accession: A34244
 C:Species: Rattus norvegicus (Norway rat)
 C:Title: 07-Sep-1993 #sequence_revision 02-Jun-1994 #text_change 28-Feb-1997
 C:Accession: A34244
 K:Magnum, M.A.; Shelton, K.D.
 J. Biol. Chem. 264: 15936-15942, 1989
 A:Title: An alternate promoter in the alkaline phosphatase gene is active in the pancreatic
 A:Reference number: A34244; M010-89480186
 A:Accession: A34244
 A:Molecule type: mRNA
 A:Residues: 1-28 (28)
 A:Cross-references: DB:M25897
 A:Note: the authors translated the codon CTG for residue 2 as Lys
 C:Superfamily: hexokinase; hexokinase homology
 C:Keywords: ATP; hydrolysis; phosphatase

 Query Match: 29.4%; Score 24; DB 2; Length 28;
 Best Local Similarity 42.9%; Pred. No. 1.4e+03;
 Matches 6; Conservative 2; Mismatches 6; Indels 0; Gaps 0;
 Q7 1 KADKATPKKVSSE 14
 I0 5 KADKATPKKVSSE 18

Search completed: May 1, 2001, 11:26:22
 Job time: 127 sec




```

1  * MEDIUM: IBM PC compatible
2  OPERATING SYSTEM: PC DOS/MS DOS
3  SOFTWARE: PATENT IN PROGRESS #1.0, Version #1.20
4  CURRENT APPLICATION DATA:
5  APPLICANT: R. 879045
6  FILING DATE: 04 APR-1996
7  CLASSIFICATION: 435
8  PRIOR APPLICATION DATA:
9  APPLICATION NUMBER: 148064/1995
10 FILING DATE: 24 MAY-1995
11 ATTORNEY/AGENT INFORMATION:
12 NAME: ORLAN, NORMAN F
13 REGISTRATION NUMBER: 24,618
14 REFERENCE/CHECK NUMBER: 7514 501 0
15 TELECOMMUNICATION INFORMATION:
16 TELEPHONE: (704) 413-4000
17 TELEFAX: (704) 413-2220
18 TELEX: 240055 OPAI UR
19 INFORMATION FOR SEQ ID NO: 2:
20 SEQUENCE CHARACTERISTICS:
21 LENGTH: 14 amino acids
22 TYPE: amino acid
23 STRANDEDNESS: single
24 FEATURE:
25 FEATURE TYPE: peptide
26 TAG/KEY: Modified site
27 LOCATION: 12
28 OTHER INFORMATION: N- (14-22) "KLF" at 21, 26."
29 US 08 540 897 2

```

Query Match 41.5% Score 34; DB 1; Length 19;
 Best local Similarity 43.8%; Pred. No. 8.2;
 Matches 7; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

```

29 1 KRLRTPKHKVSEQR 16
30 1 1 1 1 1 1 1 1
31 3 KRPETAEHLVERQ 26

```

```

1  * MEDIUM: IBM PC compatible
2  OPERATING SYSTEM: PC DOS/MS DOS
3  SOFTWARE: PATENT IN PROGRESS #1.0, Version #1.20
4  CURRENT APPLICATION DATA:
5  APPLICANT: R. 879045
6  FILING DATE: 04 APR-1996
7  CLASSIFICATION: 435
8  PRIOR APPLICATION DATA:
9  APPLICATION NUMBER: 148064/1995
10 FILING DATE: 24 MAY-1995
11 ATTORNEY/AGENT INFORMATION:
12 NAME: ORLAN, NORMAN F
13 REGISTRATION NUMBER: 24,618
14 REFERENCE/CHECK NUMBER: 7514 501 0
15 TELECOMMUNICATION INFORMATION:
16 TELEPHONE: (704) 413-4000
17 TELEFAX: (704) 413-2220
18 TELEX: 240055 OPAI UR
19 INFORMATION FOR SEQ ID NO: 2:
20 SEQUENCE CHARACTERISTICS:
21 LENGTH: 14 amino acids
22 TYPE: amino acid
23 STRANDEDNESS: single
24 FEATURE:
25 FEATURE TYPE: peptide
26 TAG/KEY: Modified site
27 LOCATION: 12
28 OTHER INFORMATION: N- (14-22) "KLF" at 21, 26."
29 US 08 540 897 2

```

Query Match 41.5% Score 34; DB 1; Length 19;
 Best local Similarity 43.8%; Pred. No. 8.2;
 Matches 7; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

```

29 1 KRLRTPKHKVSEQR 16
30 1 1 1 1 1 1 1 1
31 3 KRPETAEHLVERQ 26

```

```

1  * MEDIUM: IBM PC compatible
2  OPERATING SYSTEM: PC DOS/MS DOS
3  SOFTWARE: PATENT IN PROGRESS #1.0, Version #1.20
4  CURRENT APPLICATION DATA:
5  APPLICANT: R. 879045
6  FILING DATE: 04 APR-1996
7  CLASSIFICATION: 435
8  PRIOR APPLICATION DATA:
9  APPLICATION NUMBER: 148064/1995
10 FILING DATE: 24 MAY-1995
11 ATTORNEY/AGENT INFORMATION:
12 NAME: ORLAN, NORMAN F
13 REGISTRATION NUMBER: 24,618
14 REFERENCE/CHECK NUMBER: 7514 501 0
15 TELECOMMUNICATION INFORMATION:
16 TELEPHONE: (704) 413-4000
17 TELEFAX: (704) 413-2220
18 TELEX: 240055 OPAI UR
19 INFORMATION FOR SEQ ID NO: 2:
20 SEQUENCE CHARACTERISTICS:
21 LENGTH: 14 amino acids
22 TYPE: amino acid
23 STRANDEDNESS: single
24 FEATURE:
25 FEATURE TYPE: peptide
26 TAG/KEY: Modified site
27 LOCATION: 12
28 OTHER INFORMATION: N- (14-22) "KLF" at 21, 26."
29 US 08 540 897 2

```

```

1  * MEDIUM: IBM PC compatible
2  OPERATING SYSTEM: PC DOS/MS DOS
3  SOFTWARE: PATENT IN PROGRESS #1.0, Version #1.20
4  CURRENT APPLICATION DATA:
5  APPLICANT: R. 879045
6  FILING DATE: 04 APR-1996
7  CLASSIFICATION: 435
8  PRIOR APPLICATION DATA:
9  APPLICATION NUMBER: 148064/1995
10 FILING DATE: 24 MAY-1995
11 ATTORNEY/AGENT INFORMATION:
12 NAME: ORLAN, NORMAN F
13 REGISTRATION NUMBER: 24,618
14 REFERENCE/CHECK NUMBER: 7514 501 0
15 TELECOMMUNICATION INFORMATION:
16 TELEPHONE: (704) 413-4000
17 TELEFAX: (704) 413-2220
18 TELEX: 240055 OPAI UR
19 INFORMATION FOR SEQ ID NO: 2:
20 SEQUENCE CHARACTERISTICS:
21 LENGTH: 28 amino acids
22 TYPE: amino acid
23 TOPOLOGY: linear
24 US 08 331 394 20

```

Query Match 39.0% Score 32; DB 1; Length 28;
 Best local Similarity 42.9%; Pred. No. 27;
 Matches 6; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

```

30 3 KRLRTPKHKVSEQR 16
31 3 1 1 1 1 1 1 1
32 13 KRPETAEHLVERQ 26

```

```

1  * MEDIUM: IBM PC compatible
2  OPERATING SYSTEM: PC DOS/MS DOS
3  SOFTWARE: PATENT IN PROGRESS #1.0, Version #1.20
4  CURRENT APPLICATION DATA:
5  APPLICANT: R. 879045
6  FILING DATE: 04 APR-1996
7  CLASSIFICATION: 435
8  PRIOR APPLICATION DATA:
9  APPLICATION NUMBER: 148064/1995
10 FILING DATE: 24 MAY-1995
11 ATTORNEY/AGENT INFORMATION:
12 NAME: ORLAN, NORMAN F
13 REGISTRATION NUMBER: 24,618
14 REFERENCE/CHECK NUMBER: 7514 501 0
15 TELECOMMUNICATION INFORMATION:
16 TELEPHONE: (704) 413-4000
17 TELEFAX: (704) 413-2220
18 TELEX: 240055 OPAI UR
19 INFORMATION FOR SEQ ID NO: 2:
20 SEQUENCE CHARACTERISTICS:
21 LENGTH: 28 amino acids
22 TYPE: amino acid
23 TOPOLOGY: linear
24 US 08 331 394 20

```

Query Match 39.0% Score 32; DB 1; Length 28;
 Best local Similarity 42.9%; Pred. No. 27;
 Matches 6; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

```

30 3 KRLRTPKHKVSEQR 16
31 3 1 1 1 1 1 1 1
32 13 KRPETAEHLVERQ 26

```

```

1  * MEDIUM: IBM PC compatible
2  OPERATING SYSTEM: PC DOS/MS DOS
3  SOFTWARE: PATENT IN PROGRESS #1.0, Version #1.20
4  CURRENT APPLICATION DATA:
5  APPLICANT: R. 879045
6  FILING DATE: 04 APR-1996
7  CLASSIFICATION: 435
8  PRIOR APPLICATION DATA:
9  APPLICATION NUMBER: 148064/1995
10 FILING DATE: 24 MAY-1995
11 ATTORNEY/AGENT INFORMATION:
12 NAME: ORLAN, NORMAN F
13 REGISTRATION NUMBER: 24,618
14 REFERENCE/CHECK NUMBER: 7514 501 0
15 TELECOMMUNICATION INFORMATION:
16 TELEPHONE: (704) 413-4000
17 TELEFAX: (704) 413-2220
18 TELEX: 240055 OPAI UR
19 INFORMATION FOR SEQ ID NO: 2:
20 SEQUENCE CHARACTERISTICS:
21 LENGTH: 28 amino acids
22 TYPE: amino acid
23 TOPOLOGY: linear
24 US 08 331 394 20

```

Query Match 39.0% Score 32; DB 1; Length 28;
 Best local Similarity 42.9%; Pred. No. 27;
 Matches 6; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

```

30 3 KRLRTPKHKVSEQR 16
31 3 1 1 1 1 1 1 1
32 13 KRPETAEHLVERQ 26

```

```

RESULT 5
US 09 341 179 20
Sequence 20, Application US/09341179
Patent No. 5/41007
GENERAL INFORMATION:
APPLICANT: Genentech, David V.
APPLICANT: Potl, Mike
TITLE OF INVENTION: Tumor Necrosis Factor Receptor-Associated Factors
NUMBER OF SEQUENCES: 66
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER RELEASABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC DOS/MS-DOS
SOFTWARE: Winpatin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09341179
FILING DATE: 31 Oct 1994
CLASSIFICATION: 4.35
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/250858
FILING DATE: 05/27/1994
ATTORNEY/AGENT INFORMATION:
NAME: Drenth, Gloriet R.
REGISTRATION NUMBER: 33,055
REFERENCE/BOOKLET NUMBER: 090701
TELEPHONE: 415/225-3216
TELEFAX: 415/952-9881
TELEX: 910/471-7168
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 28 amino acids
TYPE: Amino Acid
TOPOLGY: Linear
US 09-341-179 20

Query Match 99.08% Score 42 DB 1 Length 28
Best Local Similarity 42.96% Prod. No. 27
Matches 62 Conserved 190 2 Mismatches 6 Gaps 0

27 5 POLYPEPTIDE SEQ 16
3 1 1 1 1 1
DB 13 KEEPTAPE/VEHOR 26

RESULT 7
US 09 341 179 20
Sequence 20, Application US/09341179
Patent No. 5/41007
GENERAL INFORMATION:
APPLICANT: Genentech, Inc.
TITLE OF INVENTION: Tumor Necrosis Factor Receptor-Associated Factors
NUMBER OF SEQUENCES: 66
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER RELEASABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC DOS/MS-DOS
SOFTWARE: patin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: 09/250858
FILING DATE: 05/27/1994
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/250858
FILING DATE: 27 MAY 1994
ATTORNEY/AGENT INFORMATION:
NAME: Drenth, Gloriet R.
REGISTRATION NUMBER: 33,055
REFERENCE/BOOKLET NUMBER: 090701
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-3216

```

```

RESULT 5
US 09 341 179 20
Sequence 20, Application US/09341179
Patent No. 5/41007
GENERAL INFORMATION:
APPLICANT: Genentech, David V.
APPLICANT: Potl, Mike
TITLE OF INVENTION: Tumor Necrosis Factor Receptor-Associated Factors
NUMBER OF SEQUENCES: 66
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER RELEASABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC DOS/MS-DOS
SOFTWARE: Winpatin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09341179
FILING DATE: 31 Oct 1994
CLASSIFICATION: 4.35
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/250858
FILING DATE: 05/27/1994
ATTORNEY/AGENT INFORMATION:
NAME: Drenth, Gloriet R.
REGISTRATION NUMBER: 33,055
REFERENCE/BOOKLET NUMBER: 090701
TELEPHONE: 415/225-3216
TELEFAX: 415/952-9881
TELEX: 910/471-7168
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 28 amino acids
TYPE: Amino Acid
TOPOLGY: Linear
US 09-341-179 20

Query Match 99.08% Score 42 DB 2 Length 28
Best Local Similarity 42.96% Prod. No. 27
Matches 62 Conserved 190 2 Mismatches 6 Gaps 0

27 5 POLYPEPTIDE SEQ 16
3 1 1 1 1 1
DB 13 KEEPTAPE/VEHOR 26

RESULT 7
US 09 341 179 20
Sequence 20, Application US/09341179
Patent No. 5/41007
GENERAL INFORMATION:
APPLICANT: Genentech, Inc.
TITLE OF INVENTION: Tumor Necrosis Factor Receptor-Associated Factors
NUMBER OF SEQUENCES: 66
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER RELEASABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC DOS/MS-DOS
SOFTWARE: patin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: 09/250858
FILING DATE: 05/27/1994
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/250858
FILING DATE: 27 MAY 1994
ATTORNEY/AGENT INFORMATION:
NAME: Drenth, Gloriet R.
REGISTRATION NUMBER: 33,055
REFERENCE/BOOKLET NUMBER: 090701
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-3216

```

TELEFAX: 415/952-9881
 TELEX: 910/571-7169
 INFORMATION FOR SEQ ID NO: 20:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 28 amino acids
 TYPE: amino acid
 STRAND: Linear
 PRT-US96-066 (9/20)

Query Match 47.8% Score 32, DB 5, Length 28,
 Best Local Similarity 100.0%, Ident. No. 17,
 Mismatches 0, Conservative 0, Gaps 0,

QY 8 RUDHPIKVSQAR 16
 : 1 1 1 1 1
 DQ 14 KETAPHEVER 26

RESULT 9
 PCT US96-066 54-142

Sequence 142, Application US/08548540
 Patent No. 5743731

GENERAL INFORMATION:

APPLICANT: Schatz, Peter J.

APPLICANT: Cull, Millard G.

APPLICANT: Miller, Jeff F.

APPLICANT: Stemmer, Willem P.

APPLICANT: Gates, Christian M.

TITLE OF INVENTION: Peptide Library and Screening Method

NUMBER OF SEQUENCES: 142

CORRESPONDENCE ADDRESS:

ADDRESSEE: William M. Smith

STREET: One Market Plaza, Stewart Tower, Suite 2000

CITY: San Francisco

STATE: California

COUNTRY: USA

ZIP: 94105

COMPUTER READABLE FORM:

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US 08/548540

FILING DATE: 26 OCT 1995

CLASSIFICATION: 45

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/290,641

FILING DATE: 15-AUG-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/563,321

FILING DATE: 15 OCT 1992

ADDITIONAL INFORMATION:

NAME: Smith, William M.

REGISTRATION NUMBER: 30,223

REFERENCE/SEQ ID NO: 142

TELECOMMUNICATION INFORMATION:

TELEPHONE: 415-226-2400

TELEFAX: 415-226-2422

INFORMATION FOR SEQ ID NO: 142:

SEQUENCE CHARACTERISTICS:

LENGTH: 12 amino acids

TYPE: amino acid

STRANDNESS: single

Topology: linear

MOLECULE TYPE: peptide

PRT-US96-066 132

Query Match 47.8% Score 32, DB 5, Length 12,
 Best Local Similarity 100.0%, Ident. No. 17,
 Mismatches 0, Conservative 0, Gaps 0,

QY 10 HKVSEQ 15
 : 1 1 1 1 1
 DQ 4 HKVSEQ 9

RESULT 9

PCT US96-09809-132

Sequence 132, Application PC/TUS96-09809

GENERAL INFORMATION:

APPLICANT: Schatz, Peter J.

APPLICANT: Cull, Millard G.

APPLICANT: Miller, Jeff F.

APPLICANT: Stemmer, Willem P.

APPLICANT: Gates, Christian M.

TITLE OF INVENTION: Peptide Library and Screening Method

NUMBER OF SEQUENCES: 132

CORRESPONDENCE ADDRESS:

ADDRESSEE: William M. Smith

STREET: One Market Plaza, Stewart Tower, Suite 2000

CITY: San Francisco

STATE: California

COUNTRY: USA

ZIP: 94105

COMPUTER READABLE FORM:

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US 08/290,641

FILING DATE: 15-AUG-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/563,321

FILING DATE: 15-OCT-1992

ADDITIONAL INFORMATION:

NAME: Smith, William M.

REGISTRATION NUMBER: 30,223

REFERENCE/SEQ ID NO: 142

TELECOMMUNICATION INFORMATION:

TELEPHONE: 415-226-2400

TELEFAX: 415-226-2422

INFORMATION FOR SEQ ID NO: 132:

SEQUENCE CHARACTERISTICS:

LENGTH: 12 amino acids

TYPE: amino acid

STRANDNESS: single

Topology: linear

MOLECULE TYPE: peptide

PCT US96-09809 132

Query Match 47.8% Score 31, DB 5, Length 12,

Best Local Similarity 100.0%, Ident. No. 17,

Mismatches 0, Conservative 0, Gaps 0,

QY 10 HKVSEQ 15

: 1 1 1 1 1

DQ 4 HKVSEQ 9

RESULT 10

5178861-17

Patent No. 5178861

APPLICANT: VERGARA, GILLES ROSE, ARIENS, JEROME, ARTHUR

CRUGERWING, RUTH S., HENSBERG, VICTOR R.

TITLE OF INVENTION: CROSS REACTIVE AND PROTECTIVE EPITOPES

TELEFAX: 415-494-0792

TELEX: 706141

INFORMATION FOR SEQ ID NO: 12:

SEQUENCE CHARACTERISTICS:

LENGTH: 20 amino acids

TYPE: amino acid

STRANDEDNESS: single

Topology: linear

US-09-411-859-12

Query Match

Best Local Similarity 46.6%; Score 30; DB 2; Length 20;

Mismatches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 KARDLTP 8

11 11

10 5 KALFETFP 12

RESULT 14

US-09-745-111-7

Sequence 7, Application US/06746111

Patent No. 656778

GENERAL INFORMATION:

APPLICANT: Glaxo, David

APPLICANT: Cui, Jisong

TITLE OF INVENTION: Compositions And Methods For Screening

TITLE OF INVENTION: Compounds For Anticouplant Activity

REFERENCE SEQUENCES 52

CORRESPONDENCE ADDRESS:

ADDRESSEE: Medley & Carroll, LLP

STREET: 220 Montgomery Street, Suite 2200

CITY: San Francisco

STATE: California

COUNTRY: United States of America

ZIP: 94104

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPILED: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.40

CURRENT APPLICATION DATA:

APPLICATION NUMBER: 06/746,111

FILING DATE: 06-NOV-1996

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Tomella, Diane E.

REGISTRATION NUMBER: 40,927

REFERENCE/EXCERPT NUMBER: 06 02536

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 705-8410

TELEFAX: (415) 497-8438

INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:

LENGTH: 21 amino acids

TYPE: amino acid

STRANDEDNESS: not relevant

Topology: not relevant

Molecule TYPE: protein

US-09-745-111-7

Query Match

Best Local Similarity 45.4%; Score 29; DB 4; Length 21;

Mismatches 6; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 2 ARDLTKKRVSR 16

11 11

10 5 AKETPKKLVSR 21

RESULT 15

US-09-282-351A-42

Sequence 42, Application US/09282351A

Patent No. 6194141

GENERAL INFORMATION:

APPLICANT: Paul, Aulko V.

APPLICANT: Wimmer, Eckard

APPLICANT: Kiedor, Elizabeth

TITLE OF INVENTION: INHIBITION OF PICO RNA VIRUS GPROTEIN REPLICATION BY

TITLE OF INVENTION: INTERFERENCE WITH VPG NUCLEOTID HYDRATION AND ELONGATION

FILE REFERENCE: The Research Foundation of Stonybrook

CURRENT APPLICATION NUMBER: 09/042923, 351A

CURRENT FILING DATE: 1999-03-31

NUMBER OF SEQ ID NOS: 49

SOFTWARE: Patent In Ver. 2.1

SEQ ID NO 32

LENGTH: 21

TYPE: PRT

ORGANISM: Unknown organism

FEATURES:

OTHER INFORMATION: Description of Unknown Organism: Polyomavirus

US-09-282-351A-42

Query Match

Best Local Similarity 45.4%; Score 29; DB 4; Length 21;

Mismatches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 8 PDKRVSR 16

11 11 11

10 5 PDKRVSR 17

Search completed: May 1, 2001, 11:27:43

Job time: 153 sec



CV: claimed targetting peptide.

XX Sequence: 17 AA:

Query Match 42.7% Score 45; DB 19; Length 17;

Best Local Similarity 60.00; Prod. No. 94;

Matches 6; Conservative 2; Mismatches 0; Gaps 0;

Q/ 4 EAAAGHCHG 12

11 :||| 11

DB 5 ddaqrdaqr 14

RESULT 9

W0587

ID 90587 standard; peptide: 19 AA

XX W0587

XX W0587

XX 16-JUL-1998 (first entry)

XX GPIIb/IIIa receptor ligand with thrombus imaging: GPIIb/IIIa receptor: Human:

XX Technetium-99m labeled; thrombus imaging: GPIIb/IIIa receptor: Human:

XX KW binding moiety.

XX KW Technetium-99m labeled; thrombus imaging: Specific binding:

XX KW Technetium-99m radiolabeled; GPIIb/IIIa receptor ligand.

XX KW Technetium-99m radiolabeled; GPIIb/IIIa receptor ligand.

XX KW Technetium-99m radiolabeled; GPIIb/IIIa receptor ligand.

XX KW Technetium-99m radiolabeled; GPIIb/IIIa receptor ligand.

XX KW Technetium-99m radiolabeled; GPIIb/IIIa receptor ligand.

XX KW Technetium-99m radiolabeled; GPIIb/IIIa receptor ligand.

XX KW Technetium-99m radiolabeled; GPIIb/IIIa receptor ligand.

XX KW Technetium-99m radiolabeled; GPIIb/IIIa receptor ligand.

XX KW Technetium-99m radiolabeled; GPIIb/IIIa receptor ligand.

XX KW Technetium-99m radiolabeled; GPIIb/IIIa receptor ligand.

XX KW Technetium-99m radiolabeled; GPIIb/IIIa receptor ligand.

XX KW Technetium-99m radiolabeled; GPIIb/IIIa receptor ligand.

XX KW Technetium-99m radiolabeled; GPIIb/IIIa receptor ligand.

XX KW Technetium-99m radiolabeled; GPIIb/IIIa receptor ligand.

XX KW Technetium-99m radiolabeled; GPIIb/IIIa receptor ligand.

XX KW Technetium-99m radiolabeled; GPIIb/IIIa receptor ligand.

XX KW Technetium-99m radiolabeled; GPIIb/IIIa receptor ligand.

XX KW Technetium-99m radiolabeled; GPIIb/IIIa receptor ligand.

XX KW Technetium-99m radiolabeled; GPIIb/IIIa receptor ligand.

XX KW Technetium-99m radiolabeled; GPIIb/IIIa receptor ligand.

XX KW Technetium-99m radiolabeled; GPIIb/IIIa receptor ligand.

XX KW Technetium-99m radiolabeled; GPIIb/IIIa receptor ligand.

XX KW Technetium-99m radiolabeled; GPIIb/IIIa receptor ligand.

XX KW Technetium-99m radiolabeled; GPIIb/IIIa receptor ligand.

XX KW Technetium-99m radiolabeled; GPIIb/IIIa receptor ligand.

XX KW Technetium-99m radiolabeled; GPIIb/IIIa receptor ligand.

XX KW Technetium-99m radiolabeled; GPIIb/IIIa receptor ligand.

XX KW Technetium-99m radiolabeled; GPIIb/IIIa receptor ligand.

XX KW Technetium-99m radiolabeled; GPIIb/IIIa receptor ligand.

XX KW Technetium-99m radiolabeled; GPIIb/IIIa receptor ligand.

XX KW Technetium-99m radiolabeled; GPIIb/IIIa receptor ligand.

XX KW Technetium-99m radiolabeled; GPIIb/IIIa receptor ligand.

XX KW Technetium-99m radiolabeled; GPIIb/IIIa receptor ligand.

XX KW Technetium-99m radiolabeled; GPIIb/IIIa receptor ligand.

XX KW Technetium-99m radiolabeled; GPIIb/IIIa receptor ligand.

XX KW Technetium-99m radiolabeled; GPIIb/IIIa receptor ligand.

XX KW Technetium-99m radiolabeled; GPIIb/IIIa receptor ligand.

XX KW Technetium-99m radiolabeled; GPIIb/IIIa receptor ligand.

XX KW Technetium-99m radiolabeled; GPIIb/IIIa receptor ligand.

XX KW Technetium-99m radiolabeled; GPIIb/IIIa receptor ligand.

XX KW Technetium-99m radiolabeled; GPIIb/IIIa receptor ligand.

XX KW Technetium-99m radiolabeled; GPIIb/IIIa receptor ligand.

XX KW Technetium-99m radiolabeled; GPIIb/IIIa receptor ligand.

XX KW Technetium-99m radiolabeled; GPIIb/IIIa receptor ligand.

XX KW Technetium-99m radiolabeled; GPIIb/IIIa receptor ligand.

XX KW Technetium-99m radiolabeled; GPIIb/IIIa receptor ligand.

XX KW Technetium-99m radiolabeled; GPIIb/IIIa receptor ligand.

XX KW Technetium-99m radiolabeled; GPIIb/IIIa receptor ligand.

XX KW Technetium-99m radiolabeled; GPIIb/IIIa receptor ligand.

XX KW Technetium-99m radiolabeled; GPIIb/IIIa receptor ligand.

XX KW Technetium-99m radiolabeled; GPIIb/IIIa receptor ligand.

XX KW Technetium-99m radiolabeled; GPIIb/IIIa receptor ligand.

XX KW Technetium-99m radiolabeled; GPIIb/IIIa receptor ligand.

XX KW Technetium-99m radiolabeled; GPIIb/IIIa receptor ligand.

XX KW Technetium-99m radiolabeled; GPIIb/IIIa receptor ligand.

XX KW Technetium-99m radiolabeled; GPIIb/IIIa receptor ligand.

XX KW Technetium-99m radiolabeled; GPIIb/IIIa receptor ligand.


```

NAME: CLIFF, JAMES A
REGISTRATION NUMBER: 27,075
REFERENCE/AGENT NUMBER: 4110
FILE INFORMATION:
TELEPHONE: (703) 836 6400
TELEFAX: (703) 836 2787
INFORMATION FOR SEQ ID NOS: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 24 amino acids
TYPE: amino acid
SIMILARNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHEICAL: No
US 08-461-947-2

Query Match: 41.1% Score 44; DB 4; Length 24;
Best Local Similarity 58.4%; Pred. No. 3.2;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0.

QY 1 PPGAGAGGPPG 12
   1 1 1 1 1 1
IR 5 PPGKRGAGG 16

RESULT 4
US 08-440-594-B
Sequence No. Application US/08430754-B
Patent No. 5702506
GENERAL INFORMATION:
APPLICANT: CAREY, DARRELL H.
FILE NO. INVENTION: CYTOLYTIC PEPTIDE RESTRICTED CELL-ADHESIVE AGENTS
FILE REFERENCE: CHRP-002
CURRENT APPLICATION NUMBER: 08-440-594-B
CURRENT FILING DATE: 1994-10-28
NUMBER OF SEQ ID NOS: 9
SOFTWARE: Patent Ver. 2.0
SEQ ID NO. 8
LENGTH: 24
TYPE: PKI
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US 08-440-594-B

Query Match: 41.1% Score 44; DB 4; Length 24;
Best Local Similarity 58.4%; Pred. No. 3.2;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0.

QY 1 PPGAGAGGPPG 12
   1 1 1 1 1 1
IR 5 PPGKRGAGG 16

RESULT 4
US 08-461-947-4
Sequence No. Application US/08451947
Patent No. 5702506
GENERAL INFORMATION:
APPLICANT: GENENTECH, INC.
APPLICANT: ROSENTHAL, ARON
TITLE OF INVENTION: NOVEL NEUROTROPHIC FACTOR
NUMBER OF SEQUENCES: 100
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC DOS/MS DOS
SOFTWARE: Patin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08451947
FILING DATE:
CLASSIFICATION: 514
PRIOR APPLICATION DATA:

```

```

ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC DOS/MS DOS
SOFTWARE: Patin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08451947
FILING DATE:
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/426419
FILING DATE: 19-APR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/040013
FILING DATE: 22 MAR 1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/484442
FILING DATE: 31 JAN
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/487737
FILING DATE: 1991
ATTORNEY/AGENT INFORMATION:
NAME: Torchia, Timothy E.
REGISTRATION NUMBER: 36,700
REFERENCE/AGENT NUMBER: 444F2-102
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-8674
TELEFAX: 415/262-9881
TELEX: 910231 7163
INFORMATION FOR SEQ ID NOS: 40:
SEQUENCE CHARACTERISTICS:
LENGTH: 30 amino acids
TYPE: amino acid
TOPOLOGY: linear
US 08-451-947-40

Query Match: 40.2% Score 43; DB 1; Length 40;
Best Local Similarity 80.0%; Pred. No. 5.8;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0.

QY 4 GAGGDERGV 14
   1 1 1 1 1 1
IR 12 GAGGGGGRV 21

RESULT 5
US 08-451-947-43
Sequence No. Application US/08451947
Patent No. 5702506
GENERAL INFORMATION:
APPLICANT: GENENTECH, INC.
APPLICANT: ROSENTHAL, ARON
TITLE OF INVENTION: NOVEL NEUROTROPHIC FACTOR
NUMBER OF SEQUENCES: 100
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC DOS/MS DOS
SOFTWARE: Patin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08451947
FILING DATE:
CLASSIFICATION: 514
PRIOR APPLICATION DATA:

```

1 APPLICATION NUMBER: 08/243482
 2 FILING DATE: 19 APR 1995
 3 PRIOR APPLICATION DATA:
 4 APPLICATION NUMBER: 08/243481
 5 FILING DATE: 22 MAR 1995
 6 PRIOR APPLICATION DATA:
 7 APPLICATION NUMBER: 07/643482
 8 FILING DATE: 31 JAN 1995
 9 PRIOR APPLICATION DATA:
 10 APPLICATION NUMBER: 07/643482
 11 FILING DATE: 31 JAN 1995
 12 APPLICATION NUMBER: 07/643482
 13 FILING DATE: 31 JAN 1995
 14 NAME: Torchia, Timothy E.
 15 REGISTERED AGENT: Timothy E. Torchia
 16 TELECOMMUNICATION INFORMATION:
 17 TELEPHONE: 415/225-8674
 18 TELEFAX: 415/225-8674
 19 TELEFAX: 415/225-8674
 20 TELEFAX: 415/225-8674
 21 INFORMATION FOR SEQ ID NO: 43:
 22 SEQUENCE CHARACTERISTICS:
 23 LENGTH: 40 amino acids
 24 TYPE: Amino Acid
 25 TOPOLOGY: Linear
 26 US-08-424-826A-43

Query Match: 40.2% Score: 43; DB: 1; Length: 40;
 Best Local Similarity: 80.0%; Pred. No: 5.8;
 Matches: 8; Conserved: 0; Mismatches: 2; Indels: 0; Gaps: 0;

Q7 4 GAGGAGGAGV 13
 11 11111
 14 11 GAGGAGGAGV 20

RESULT 7
 US-08-424-826A-43
 1 Sequence 43, Application US/08424826A
 2 Patent No. 5830858
 3 GENERAL INFORMATION:
 4 APPLICANT: Rosenthal, Aaron
 5 TITLE OF INVENTION: SEVEL REPRODUCTION FACTOR
 6 NUMBER OF SEQUENCES: 98
 7 CORRESPONDENCE ADDRESS:
 8 ADDRESSEE: Genentech, Inc.
 9 STREET: 460 Point San Bruno Blvd
 10 CITY: South San Francisco
 11 STATE: California
 12 COUNTRY: USA
 13 ZIP: 94080
 14 COMPUTER READABLE FORM:
 15 MEDIUM TYPE: 5.5 inch, 1.44 MB floppy disk
 16 OPERATING SYSTEM: PC DOS/MS DOS
 17 SOFTWARE: WinData (Genentech)
 18 APPLICATION NUMBER: 08/243482
 19 FILING DATE: 31 JAN 1995
 20 CLASSIFICATION: 514
 21 PRIOR APPLICATION DATA:
 22 APPLICATION NUMBER: 07/643482
 23 FILING DATE: 31 JAN 1995
 24 APPLICATION NUMBER: 07/643482
 25 FILING DATE: 31 JAN 1995
 26 APPLICATION NUMBER: 07/643482
 27 FILING DATE: 31 JAN 1995
 28 NAME: Torchia, Timothy E.
 29 REGISTERED AGENT: Timothy E. Torchia
 30 TELEPHONE: 415/225-8674
 31 TELEFAX: 415/225-8674
 32 TELEFAX: 415/225-8674
 33 INFORMATION FOR SEQ ID NO: 43:
 34 SEQUENCE CHARACTERISTICS:
 35 LENGTH: 40 amino acids
 36 TYPE: Amino Acid
 37 TOPOLOGY: Linear
 38 US-08-424-826A-43

1 APPLICATION NUMBER: 08/243482
 2 FILING DATE: 19 APR 1995
 3 PRIOR APPLICATION DATA:
 4 APPLICATION NUMBER: 08/243481
 5 FILING DATE: 22 MAR 1995
 6 PRIOR APPLICATION DATA:
 7 APPLICATION NUMBER: 07/643482
 8 FILING DATE: 31 JAN 1995
 9 PRIOR APPLICATION DATA:
 10 APPLICATION NUMBER: 07/643482
 11 FILING DATE: 31 JAN 1995
 12 APPLICATION NUMBER: 07/643482
 13 FILING DATE: 31 JAN 1995
 14 NAME: Torchia, Timothy E.
 15 REGISTERED AGENT: Timothy E. Torchia
 16 TELECOMMUNICATION INFORMATION:
 17 TELEPHONE: 415/225-8674
 18 TELEFAX: 415/225-8674
 19 TELEFAX: 415/225-8674
 20 TELEFAX: 415/225-8674
 21 INFORMATION FOR SEQ ID NO: 40:
 22 SEQUENCE CHARACTERISTICS:
 23 LENGTH: 30 amino acids
 24 TYPE: Amino Acid
 25 TOPOLOGY: Linear
 26 US-08-424-826A-40

Query Match: 40.2% Score: 43; DB: 2; Length: 40;
 Best Local Similarity: 80.0%; Pred. No: 5.8;
 Matches: 8; Conserved: 0; Mismatches: 2; Indels: 0; Gaps: 0;

Q7 4 GAGGAGGAGV 13
 11 11111
 14 12 GAGGAGGAGV 21

RESULT 7
 US-08-424-826A-43
 1 Sequence 43, Application US/08424826A
 2 Patent No. 5830858
 3 GENERAL INFORMATION:
 4 APPLICANT: Rosenthal, Aaron
 5 TITLE OF INVENTION: SEVEL REPRODUCTION FACTOR
 6 NUMBER OF SEQUENCES: 98
 7 CORRESPONDENCE ADDRESS:
 8 ADDRESSEE: Genentech, Inc.
 9 STREET: 460 Point San Bruno Blvd
 10 CITY: South San Francisco
 11 STATE: California
 12 COUNTRY: USA
 13 ZIP: 94080
 14 COMPUTER READABLE FORM:
 15 MEDIUM TYPE: 5.5 inch, 1.44 MB floppy disk
 16 OPERATING SYSTEM: PC DOS/MS DOS
 17 SOFTWARE: WinData (Genentech)
 18 APPLICATION NUMBER: 08/243482
 19 FILING DATE: 31 JAN 1995
 20 CLASSIFICATION: 514
 21 PRIOR APPLICATION DATA:
 22 APPLICATION NUMBER: 07/643482
 23 FILING DATE: 31 JAN 1995
 24 APPLICATION NUMBER: 07/643482
 25 FILING DATE: 31 JAN 1995
 26 APPLICATION NUMBER: 07/643482
 27 FILING DATE: 31 JAN 1995
 28 NAME: Torchia, Timothy E.
 29 REGISTERED AGENT: Timothy E. Torchia
 30 TELEPHONE: 415/225-8674
 31 TELEFAX: 415/225-8674
 32 TELEFAX: 415/225-8674
 33 INFORMATION FOR SEQ ID NO: 43:
 34 SEQUENCE CHARACTERISTICS:
 35 LENGTH: 40 amino acids
 36 TYPE: Amino Acid
 37 TOPOLOGY: Linear
 38 US-08-424-826A-43

Query Match: 40.2% Score: 43; DB: 2; Length: 40;
 Best Local Similarity: 80.0%; Pred. No: 5.8;
 Matches: 8; Conserved: 0; Mismatches: 2; Indels: 0; Gaps: 0;

```

QY 4 GAGGAGGCGV 13
14 11 GAGGAGGCGV 20

RESULT 10
US-08-928-694-40
? Sequence 40, Application US/08/928694
? Patent No. 6037420
? GENERAL INFORMATION:
? APPLICANT: ROSENTHAL, ARNON
? TITLE OF INVENTION: NOVEL NEUROTROPHIC FACTOR
? NUMBER OF SEQUENCES: 100
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Genentech, Inc.
? STREET: 1 DNA Way
? CITY: South San Francisco
? STATE: California
? COUNTRY: USA
? ZIP: 94080
? COMPUTER READABLE FORM:
? MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: WinPatIn (Genentech)
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: 08/0928694
? FILING DATE: 12-Sep-1997
? CLASSIFICATION: 424
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: 08/451947
? FILING DATE: 26-MAY-1995
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: 08/426419
? FILING DATE: 19-APR-1995
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: 08/040013
? FILING DATE: 22-MAR-1993
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: 07/648482
? FILING DATE: 31-JAN
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: 07/587707
? FILING DATE: 1991
? ATTORNEY/AGENT INFORMATION:
? NAME: Torchia, PhD., Timothy E.
? REGISTRATION NUMBER: 36,700
? REFERENCE/FILE NUMBER: 1056612C11021
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 650/225-9881
? TELEFAX: 650/225-9874
? INFORMATION FOR SEQ ID NO: 40:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 40 amino acids
? TYPE: Amino Acid
? TOPOLOGY: Linear
US-08-928-694-40

```

```

Query Match 40.2% Score 43 23 3 Length 30:
Best Local Similarity 80.0% Pred. No. 5.8:
Matches 9 Conservative 0 Mismatches 2 Indels 0 Gaps 0

QY 4 GAGGAGGCGV 13
14 12 GAGGAGGCGV 21

RESULT 9
US-08-928-694-43
? Sequence 43, Application US/08/928694
? Patent No. 6037420

```

```

? GENERAL INFORMATION:
? APPLICANT: ROSENTHAL, ARNON
? TITLE OF INVENTION: NOVEL NEUROTROPHIC FACTOR
? NUMBER OF SEQUENCES: 100
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Genentech, Inc.
? STREET: 1 DNA Way
? CITY: South San Francisco
? STATE: California
? COUNTRY: USA
? ZIP: 94080
? COMPUTER READABLE FORM:
? MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: WinPatIn (Genentech)
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: 08/0928694
? FILING DATE: 12-Sep-1997
? CLASSIFICATION: 424
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: 08/451947
? FILING DATE: 26-MAY-1995
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: 08/426419
? FILING DATE: 19-APR-1995
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: 08/040013
? FILING DATE: 22-MAR-1993
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: 07/648482
? FILING DATE: 31-JAN
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: 07/587707
? FILING DATE: 1991
? ATTORNEY/AGENT INFORMATION:
? NAME: Torchia, PhD., Timothy E.
? REGISTRATION NUMBER: 36,700
? REFERENCE/FILE NUMBER: 1056612C11021
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 650/225-9881
? TELEFAX: 650/225-9874
? INFORMATION FOR SEQ ID NO: 43:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 40 amino acids
? TYPE: Amino Acid
? TOPOLOGY: Linear
US-08-928-694-43

Query Match 40.2% Score 43 23 3 Length 30:
Best Local Similarity 80.0% Pred. No. 5.8:
Matches 9 Conservative 0 Mismatches 2 Indels 0 Gaps 0

QY 4 GAGGAGGCGV 13
14 11 GAGGAGGCGV 20

RESULT 10
PCT-US91-06950-40
? Sequence 40, Application PCT/US9106950
? GENERAL INFORMATION:
? APPLICANT: ROSENTHAL, ARNON
? TITLE OF INVENTION: NOVEL NEUROTROPHIC FACTOR
? NUMBER OF SEQUENCES: 100
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Genentech, Inc.
? STREET: 460 Point San Bruno Blvd
? CITY: South San Francisco
? STATE: California
? COUNTRY: USA

```

```

1  ZIP: 94080
2  COMPUTER RELEASABLE FORM:
3  MEDIUM TYPE: 5.25 inch, 400 Kb floppy disk
4  COMPUTER: IBM PC compatible
5  OPERATING SYSTEM: PC DOS/MS-DOS
6  SOFTWARE: Path (Genentech)
7  CURRENT APPLICATION DATA:
8  APPLICATION NUMBER: 01/587797
9  FILING DATE: 19910924
10 CLASSIFICATION: 436
11 PRIOR APPLICATION DATA:
12 APPLICATION NUMBER: 07/648482
13 PRIOR APPLICATION DATA:
14 APPLICATION NUMBER: 01/587707
15 ATTORNEY/AGENT INFORMATION:
16 NAME: Boushey, Max D.
17 REGISTRATION NUMBER: 27,043
18 REFERENCE NUMBER: 666P1
19 TELECOMMUNICATION INFORMATION:
20 TELEPHONE: 415/255-1994
21 TELEFAX: 415/255-9881
22 TELEX: 415/255-1994
23 INFORMATION FOR SEQ ID NO: 40:
24 SEQUENCE CHARACTERISTICS:
25 LENGTH: 40 amino acids
26 TYPE: AMINO ACID
27 TOPOLOGY: linear
28 US91-06950-4

```

```

Query Match: 40.2% Score 43 DB 5: Length 40:
Best Local Similarity: 80.0% Ident. No. 5.8:
Matches 6: Conservation 0: Mismatches 2: Gaps 0:

```

```

29 4 GAGGQGRGV 13

```

```

DB 11 11111
12 GAGGQGRGV 21

```

```

RESULT 11

```

```

1 Sequence 43, Application patent 436
2 GENERAL INFORMATION:
3 APPLICANT: GENENTECH, INC.
4 APPLICANT: GENENTECH, INC.
5 TITLE OF INVENTION: N-AMYL NHEER-IP-EPH-IP-FA-FA-FA
6 NUMBER OF SEQUENCES: 169
7 CORRESPONDENCE ADDRESS:
8 ADDRESSEE: Genentech, Inc.
9 STREET: 460 Point San Bruno Blvd
10 CITY: South San Francisco
11 STATE: California
12 COUNTRY: USA
13 ZIP: 94080
14 COMPUTER RELEASABLE FORM:
15 MEDIUM TYPE: 5.25 inch, 400 Kb floppy disk
16 COMPUTER: IBM PC compatible
17 OPERATING SYSTEM: PC DOS/MS-DOS
18 SOFTWARE: Path (Genentech)
19 CURRENT APPLICATION DATA:
20 APPLICATION NUMBER: 01/587797
21 FILING DATE: 19910924
22 CLASSIFICATION: 436
23 PRIOR APPLICATION DATA:
24 APPLICATION NUMBER: 07/648482
25 PRIOR APPLICATION DATA:
26 APPLICATION NUMBER: 01/587797
27 ATTORNEY/AGENT INFORMATION:
28 NAME: Boushey, Max D.
29 REGISTRATION NUMBER: 27,043
30 TELECOMMUNICATION INFORMATION:
31 TELEPHONE: 415/255-1994

```

```

1  TELEFAX: 415/255-9881
2  TELEX: 415/255-1994
3  INFORMATION FOR SEQ ID NO: 43:
4  SEQUENCE CHARACTERISTICS:
5  LENGTH: 40 amino acids
6  TYPE: AMINO ACID
7  TOPOLOGY: linear
8  US91-06950-4

```

```

Query Match: 40.2% Score 43 DB 5: Length 40:
Best Local Similarity: 80.0% Ident. No. 5.8:
Matches 6: Conservation 0: Mismatches 2: Gaps 0:

```

```

QY 4 GAGGQGRGV 13

```

```

DB 11 11111
12 GAGGQGRGV 20

```

```

RESULT 12

```

```

1 Sequence 43, Application patent 436
2 Patent No. 614824
3 GENERAL INFORMATION:
4 APPLICANT: HIR SHAVIT, RACHEL
5 TITLE OF INVENTION: N-AMYL NHEER-IP-EPH-IP-FA-FA-FA
6 NUMBER OF SEQUENCES: 4
7 CORRESPONDENCE ADDRESS:
8 ADDRESSEE: OLIFF & BERTRIDGE, PLLC
9 STREET: P.O. Box 19928
10 CITY: ALEXANDRIA
11 STATE: VA
12 COUNTRY: USA
13 ZIP: 22320
14 COMPUTER RELEASABLE FORM:
15 MEDIUM TYPE: floppy disk
16 COMPUTER: IBM PC compatible
17 OPERATING SYSTEM: PC DOS/MS-DOS
18 SOFTWARE: Patent Release #1.0, Version #1.0
19 CURRENT APPLICATION DATA:
20 APPLICATION NUMBER: 05/57981,088
21 FILING DATE: 27 JAN 1998
22 CLASSIFICATION: 514
23 ATTORNEY/AGENT INFORMATION:
24 NAME: OLIFF, JAMES A
25 REGISTRATION NUMBER: 27,076
26 REFERENCE NUMBER: JAS 40455
27 TELECOMMUNICATION INFORMATION:
28 TELEPHONE: (703) 836-6400
29 TELEFAX: (703) 836-2787
30 INFORMATION FOR SEQ ID NO: 4:
31 SEQUENCE CHARACTERISTICS:
32 LENGTH: 14 amino acids
33 TYPE: amino acid
34 STRANDEDNESS: single
35 TOPOLOGY: linear
36 MOLECULE TYPE: peptide
37 HYDROPHILIC: No
38 US-08-981-088 4

```

```

Query Match: 45.5% Score 58 DB 4: Length 14:
Best Local Similarity: 60.0% Ident. No. 1.5:
Matches 6: Conservation 1: Mismatches 5: Gaps 0:

```

```

QY 1 PPEGKRGVAC 10

```

```

DB 5 PPEGKRGVAC 14

```

```

RESULT 13

```

```

US-08-482-880-11

```

Sequence 11, Application us-09-341-829a-5
Patent No. 5746122
GENERAL INFORMATION:
APPLICANT: Dean, Richard I.
TITLE OF INVENTION: Technetium 99m Labeled Peptides for
TITLE OF INVENTION: Technetium 99m Labeled Peptides for
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: Banner & Allogretti, Ltd.
STREET: 10 South Wacker Drive Suite 3000
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: 05/09,482,880
FILING DATE: 07-JUN-1995
CLASSIFICATION: 424
AUTHOR/AGENT INFORMATION:
NAME: No. 584926000, Kevin E
REGISTRATION NUMBER: 35,403
REFERENCE/CKET NUMBER: 92,216-M
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-715-1000
TELEFAX: 312-715-1234
TELEX: 910 221-5317
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 16 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-341-829a-5

Query Match: 42.7% Score 35, DB 1, Length 16,
Best Local Similarity 60.0%, Pred. No. 40;
Matches 6; Conservative 2; Mismatches 2; Indels 0;
QY 3 DGGKGDGGRG 12
DB 5 DGGKGDGGRG 14

RESULT 14
US-09-341-829a-5
Sequence 11, Application us-09-341-829a-5
Patent No. 5746122
GENERAL INFORMATION:
APPLICANT: Dean, Richard I.
TITLE OF INVENTION: Technetium 99m Labeled Peptides for
TITLE OF INVENTION: Technetium 99m Labeled Peptides for
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: Banner & Allogretti, Ltd.
STREET: 10 South Wacker Drive Suite 3000
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:

Application Number: 05/09,482,880
FILING DATE: 07-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 05/07,936,752
FILING DATE: 21-MAY-1992
AUTHOR/AGENT INFORMATION:
NAME: No. 584926000, Kevin E
REGISTRATION NUMBER: 35,403
REFERENCE/CKET NUMBER: 92,216-M
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-715-1000
TELEFAX: 312-715-1234
TELEX: 910 221-5317
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 16 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-341-829a-5

Query Match: 42.7% Score 35, DB 2, Length 16,
Best Local Similarity 60.0%, Pred. No. 40;
Matches 6; Conservative 2; Mismatches 2; Indels 0;
QY 3 DGGKGDGGRG 12
DB 5 DGGKGDGGRG 14

RESULT 15
US-09-475-041-11
Sequence 11, Application US/0475041
Patent No. 5879658
GENERAL INFORMATION:
APPLICANT: Dean, Richard I.
TITLE OF INVENTION: Technetium 99m Labeled Peptides for
TITLE OF INVENTION: Technetium 99m Labeled Peptides for
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: Banner & Allogretti, Ltd.
STREET: 10 South Wacker Drive Suite 3000
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: 05/09,475,041
FILING DATE: 07-JUN-1995
CLASSIFICATION: 424
AUTHOR/AGENT INFORMATION:
NAME: No. 587965800, Kevin E
REGISTRATION NUMBER: 35,403
REFERENCE/CKET NUMBER: 92,216-M
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312 715 1000
TELEFAX: 312-715-1234
TELEX: 910-221-5317
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 16 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-475-041-11

Query Match: 42.74; Score: 0; DB: 2; Length: 16;
Best Local Similarity: 60.00; Pct: 40;
Matches: 0; Conserved: 2; Mismatches: 0; Gaps: 0;

0; 0; 0; 0; 0; 0; 0; 0; 0; 0;
10; 0; 0; 0; 0; 0; 0; 0; 0; 0;

Search completed: May 1, 2001, 11:27:42
Total time: 15.0 sec




```

RESULT 8
TRYPTOPHAN
ID TRYP_TROAI STANDARD PRI: 21 AA.
AC P11241;
DT 01-FEB-1994 (Rel. 29, Created)
DT 01-FEB-1994 (Rel. 29, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE TRYPSIN PRECURSOR (EC 3.4.21.4) (FRAGMENT)
OS Protophytus actinophilus (Marbled lungfish)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CA Dipnoi; Lepidosireniformes; Protopterygia; Protopterus.
CX NCBI_TaxID:7886;
RN 11;
RP SEQUENCE.
RF TISSUE:Pancreas;
RA Brimacombe M.A., Iye K.W., Beck J.R., Nigam S.K., Walsh K.A.,
  "Comparison of the amino terminal sequences of bovine, dogfish, and
  lungfish trypsins.",
  FEBS Lett. 14:222-224(1971).
CC 1: CATALYTIC ACTIVITY: PREFERENTIAL CLEAVAGE: ARG-, LYS-,
  2: SUBCELLULAR LOCATION: EXTRACELLULAR.
CC 3: SIMILARITY: 66% IDENTITY TO TRYPSIN FAMILY 51, A.A.A. KNOWN AS THE
  TRYPSIN FAMILY
DT PIR: A27119; A27119.
DR BSSP: P07288; LPPA.
DR MEROPS: S01151;
DR InterPro: IP001264;
DR PROSITE: PS00144; TRYPSIN_HUS; PARTIAL.
DR PROSITE: PS00145; TRYPSIN_SIB; PARTIAL.
KW Hydrolyase; Serine protease; Zymogen.
FI PR00P 1 7 ACTIVATION PEPTIDE.
FI CHAIN 2 21
FI R241LK 23 21
SI SEQUENCE 21 AA: 2454 MW. 9862PUL110729: CR64.

Query Match 28.8%; Score 24; DB 1; Length 21;
Best local similarity 80.0%; Pred. No. 4,6,0,0,2;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 8 PR00P 12
ID 11
ID 15 PR00P 19

RESULT 9
PSAL_PEA
ID PSAL_PEA STANDARD PRI: 13 AA.
AC P11229;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE PROTEIN SYSTEM 1 REACTION CENTRE SUBUNIT 1X (PSI 1) (FRAGMENT).
GN PSAL.
OS Pisum sativum (Garden pea).
OC Chloropast.
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta.
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids 1;
OC Fabales; Fabaceae; Papilionoideae; Pisum.
CX RIV_TaxID:6086;
RN 11;
RP SEQUENCE.
RF MEDLINE 98242987; PubMed 2485954;
RA French M., Birnstiel M., Dreyer A., Dreyer C., Inoue Y.,
  "Polypeptide composition of higher plant photosystem I complex.
  Identification of psal1, psal2 and psal3 gene products.",
  FEBS Lett. 26:274-278(1990).
CC 1: FUNCTION: MAY HELP IN THE ORGANIZATION OF THE PSAL AND PSAP
  SUBUNITS.
CC 2: SIMILARITY: 66% IDENTITY TO THE PSAL FAMILY.
DR PIR: S09743; S09743.
KW Chloroplast; Photosystem 1; Photosynthesis; Transmembrane.

```

```

ET TRANSFER 6 13
ET NONTER 13 13
SQ SEQUENCE 13 AA: 1516 MW. 9E2E45D11FDE3B41 CR664;

Query Match 26.8%; Score 22; DB 1; Length 13;
Best local similarity 60.0%; Pred. No. 4,1,0,0,2;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 R01KT 7
ID 11
ID 2 R01KT 6

RESULT 10
H087_H08P
ID H087_H08P STANDARD PRI: 28 AA.
AC P02864;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 01-NOV-1990 (Rel. 16, Last annotation update)
DE C SERGIN (FRAGMENT).
CC 1: Bacterium Sporichthya (Wheat Y7).
CC 2: Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
  Magnoliophyta; Liliopsida; Poales; Poaceae; Poaceae; Triticeae;
  Hordeum.
CX NCBI_TaxID:77009;
RN 11;
RP SEQUENCE.
RA Shewry P.R., Autran J.-C., Nimmo C.C., Lew E.J., Le... Kasarda D.D.,
  "N-terminal amino acid sequence homology of storage protein components
  from barley and a diploid wheat.",
  Nature. 286:520-522(1980).
CC 1: TRYPSIN-LIKE SERINE PROTEASE;
CC 2: TRYPSIN-LIKE SERINE PROTEASE;
DR PIR: A03355; A03355.
KW Seed storage protein; Multigene family.
IT NONTER 28 28
SI SEQUENCE 28 AA: 6133 MW. 80BA2BF244775AA CR664;

Query Match 26.8%; Score 22; DB 1; Length 28;
Best local similarity 21.4%; Pred. No. 9,4,0,0,2;
Matches 3; Conservative 7; Mismatches 4; Indels 0; Gaps 0;

QY 2 APMLETPKHVSEQ 15
ID 11
ID 7 S0ELQSPQSYLQQ 20

RESULT 11
V09_S0V4
ID V09_S0V4 STANDARD PRI: 28 AA.
AC P11241;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 01-JUL-1989 (Rel. 11, Last annotation update)
DE SERP 5 PEPTIDE.
GN V09.
OS Spiroplasma virus 4 (SPV4).
OC Viruses; ssDNA viruses; Microviridae; Spiromicrovirinae.
CX NCBI_TaxID:10855;
RN 11;
RP SEQUENCE FROM N.A.
RA Mendicino 99022909; PubMed 2922658;
RA Renaudin J., Pascari M. C., Rose J.M.,
  "Spiroplasma virus 4, nucleotide sequence of the viral RNA,
  regulatory signals, and proposed genome organization.",
  J. Bacteriol. 169:4950-4961(1987).
CC 1: This virus' RNA entry is available. It is produced through a collation of
  between the Swiss Institute of Bioinformatics and the EMBL Outstation
  at the European Bioinformatics Institute. There are no restrictions on its

```

use by non-profit institutions as long as its content is in no way modified and this statement is not removed, reprinted, or otherwise used by any other entity, requires a license agreement (see <http://www.elsevier.com/locate/elsevier/permissions>) or send an email to permissions@elsevier.com.

FT N_NLTER 14 14
SQ SEQUENCE 14 AA: 175% MW: 655836601670256 CRC64;

Query Match: 25.6%; Score 21; DB 1; Length 14;
Best Local Similarity: 42.9%; Pred. No. 8.9e-02;
Matches: 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

CV 4 LRTPKH 10
LE 7 EYRAPHK 14

RESULT 15

HEXADENOV
11 HEXADENOV STANDARD; PRT; 18 AA.
AC P45985;
DI 01-JUN-1994 (Ref. 29, Created)
DT 01-JUN-1994 (Ref. 29, Last sequence update)
DE 01-NOV-1997 (Ref. 45, Last annotation update)
HE HEXON PROTEIN (LATE PROTEIN 2) (FRAGMENT).
IN F11.

OS Canine adenovirus type 1 (strain Utrecht).
CA Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.
X NRI_taxid 36364;
RN 11;

RP SEQUENCE FROM N.A.
RX MEDLINE:94043182; PubMed 1414543;
RA Cat P.; Weber J.M.;
RI "Nucleotide and deduced amino acid sequence of the canine adenovirus
type 1 proteinase".
RL Virus Genes 6:407-412(1992).

CC 1- JUN1994; THIS PROTEIN IS ONE OF THE STRUCTURAL PROTEINS IN THE
VIRAL CAPSID AND IS SYNTHESIZED DURING LATE INFECTION.
C1 1- SUBUNIT; HOMOTRIMER (BY SIMILARITY).

CC THIS SWISS-PROT entry is copyrighted. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
enterprises requires a license agreement. Please apply to the EMBL outstation
or send an email to license@isb.ch).

CC
18 EMRL: M72715; AAA2528.1; ..
19 P3K: A48550; A48550.

19 HSSP: P04277; IDHX.

RW Coat protein: Hexon protein: Late protein.

FT N_NLTER 1 1

SQ SEQUENCE 18 AA: 1949 MW: 19922AE07A120962 CRC64;

Query Match: 25.6%; Score 21; DB 1; Length 18;
Best Local Similarity: 100.0%; Pred. No. 8.9e-02;
Matches: 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CV 5 LRTPKH 6
LE 7 LRTPKH 10

Search completed: May 1, 2001, 11:25:52
Job time: 102 sec

DE Q9VQV9 (TREMBL) 15, Last annotation update)
 DE ANKON EX-HAMMER AIR 4 PROTEIN (FRAGMENT)
 OS Gallus gallus (Chicken)
 EC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Actinopteria; Aves; Neognathae; Caprimulgidae; Phasianidae; Phasianinae;
 OC Gallus
 CX NP_134819.9(21)
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE: 95288119; PubMed 7408971
 FA Cox K.H.; Cox J.V.
 RI "Virus-like particles associated with peroxisome dysfunction NH(2)-terminal
 RI cytoplasmic domains."
 RL Am. J. Physiol. 268:0-0(0)
 DR EMBL: S76800; AAC14248.1 (1)
 FI N_N_TERMIN 1
 SQ SEQUENCE 28 AA; 2776 MW; 849661R34E1P959F CDS64;

Query Match 66.7% Score 24; DB 14; Length 28;
 Best Local Similarity 83.4% Pred. No. 2e+02;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Q7 1 KQAGR 6
 DE 1 KQAGR 6
 RESULT 4
 Q9VQV9
 ID Q9VQV9 PRELIMINARY; PRT: 20 AA.
 AC Q9VQV9
 DI 01-MAY-2000 (TREMBL) 14, Created)
 DI 01-MAY-2000 (TREMBL) 14, Last sequence update)
 DI 01-MAY-2000 (TREMBL) 15, Last annotation update)
 DE ALPHA CLASS SUBSTITUTION 2 (PRT 2.5e+1.12)
 DE (FRAGMENT)
 OS Rattus sp.
 EC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Rodentia; Muridae; Murinae; Rattus
 CX NP_134819.10(118)
 RN [1]
 RP SEQUENCE
 RX MEDLINE: 96045981; PubMed 7485987;
 FA Ewing P.; Beckaver L.; Jullier J.
 RI "Electrospray ionization-mass spectrometry as a tool for
 RI characterization of glutathione S-transferase isozymes."
 RL Anal. Biochem. 229:404-412(1995)
 DR BSSP: P24472; IGDK
 SQ SEQUENCE 20 AA; 2412 MW; 949998E22040B246 CDS64.

Query Match 63.9% Score 24; DB 14; Length 20;
 Best Local Similarity 66.7% Pred. No. 2.2e+02;
 Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Q7 2 EAGRM 7
 DE 10 Q3GRM 15
 RESULT 4
 Q9VQV9
 ID Q9VQV9 PRELIMINARY; PRT: 21 AA.
 AC Q9VQV9
 DI 01-MAY-2000 (TREMBL) 14, Created)
 DI 01-MAY-2000 (TREMBL) 14, Last sequence update)
 DI 01-MAY-2000 (TREMBL) 15, Last annotation update)
 DE ALPHA CLASS SUBSTITUTION 2 (PRT 2.5e+1.12)
 DE (FRAGMENT)
 OS Rattus sp.
 EC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Rodentia; Sciurognathi; Muridae; Murinae; Rattus

CX NP_134819.10(118)
 RN [1]
 RP SEQUENCE
 RX MEDLINE: 96045981; PubMed 7485987;
 FA Ewing P.; Beckaver L.; Jullier J.
 RI "Electrospray ionization-mass spectrometry as a tool for
 RI characterization of glutathione S-transferase isozymes."
 RL Anal. Biochem. 229:404-412(1995)
 DR BSSP: P24472; IGDK
 SQ SEQUENCE 21 AA; 2675 MW; AFA6DA961E2472C CDS64;

Query Match 63.9% Score 24; DB 14; Length 21;
 Best Local Similarity 66.7% Pred. No. 2.4e+02;
 Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Q7 2 EAGRM 7
 DE 11 Q3GRM 16
 RESULT 5
 Q36227
 ID Q36227 PRELIMINARY; PRT: 27 AA.
 AC Q36227
 DI 01-JAN-1998 (TREMBL) 05, Created)
 DI 01-JAN-1998 (TREMBL) 05, Last sequence update)
 DI 01-NOV-1998 (TREMBL) 08, Last annotation update)
 DE ENVELOPE GLYCOPROTEIN (FRAGMENT)
 DE ENV
 OS Human immunodeficiency virus type 1;
 OS Viruses; Retroviridae; Lentivirinae
 CX NP_134819.11(676)
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE: 96061399; PubMed 9443234;
 FA Nelson J.A.; Fiscus S.A.; Swanson R.
 RI "Neutralizing activity of the human immunodeficiency virus type 1 V
 RI region characterized by using a latex agglutination assay."
 RL J. Virol. 71:8750-8758(1997)
 DR EMBL: AF001804; AAC58485.1 (1)
 KW Envelope protein
 FT KQAGR 1
 FT KQAGR 27
 SQ SEQUENCE 27 AA; 2744 MW; 1274B04E89FEA60 CDS64;

Query Match 63.9% Score 24; DB 14; Length 27;
 Best Local Similarity 83.4% Pred. No. 3.1e+02;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Q7 1 KQAGR 6
 DE 9 K3AGR 14
 RESULT 6
 Q9VQV9
 ID Q9VQV9 PRELIMINARY; PRT: 28 AA.
 AC Q9VQV9
 DI 01-MAY-2000 (TREMBL) 14, Created)
 DI 01-MAY-2000 (TREMBL) 14, Last sequence update)
 DI 01-MAY-2000 (TREMBL) 15, Last annotation update)
 DE GLUTATHIONE S-TRANSFERASE SUBUNIT YX, GST SUBUNIT YX
 OS Rattus sp.
 EC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Rodentia; Sciurognathi; Muridae; Murinae; Rattus
 CX NP_134819.10(118)
 RN [1]
 RP SEQUENCE
 RX MEDLINE: 92233912; PubMed 1567476;
 RA Igarashi T.; Tsuchiya T.; Shikata Y.; Sakami F.; Iadaya M.; Horie T.


```

RA  Satchi L.
R1  "Biochemical aspects of a unique glutathione S-transferase subunit
R1  YZ in the liver cytosol from rats with hereditary hyperbilirubinemia.
R1  Comparison with rat total liver transferase subunit YHonus."
R1  Biochem. J. 283:607-611(1992).
R1  BSSD: 124472: 100K.
R1  UNIPROT: P60002:1.
R1  FEAM: 160043: 081: 1.
R1  SEQUENCE: 28 AA: 412 BW: 9297/9297/26 CR64.

Query Match: 61.1% Score: 23; DB: 11; Length: 28;
Best Local Similarity: 61.1%; Pred. No. 3,4e+02;
Matches: 4; Conservative: 1; Mismatches: 1; Indels: 0; Gaps: 0;

27 2 EAGRM 7
ID 11111
AC 10 EAGRM 7
PRELIMINARY: PRE: 25 AA.
D1 01 MAY 2000 (EMBLrel. 13, Created)
D1 01 OCT 2000 (EMBLrel. 15, Last sequence update)
D1 01 OCT 2000 (EMBLrel. 15, Last annotation update)
D1 01 OCT 2000 (EMBLrel. 15, Last annotation update)
DE GLUTATHIONE TRANSFERASE A3 SUBUNIT (FRAGMENT).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Murinae; Rattus.
OX NCBI_TaxID: 10116.
RN 111
SEQUENCE:
RX MEDLINE 96092035; PubMed 8529022;
RA Lindor D., Freund R., Kadenbach B.
R1 "Specific expression of cytochrome c oxidase isoenzymes."
R1 Comp. Biochem. Physiol. 112B:461-469(1995).
R1 NON_TER 1
R1 NON_TER 20
R1 SEQUENCE 28 AA: 426 BW: 9297/9297/26 CR64.

Query Match: 61.1% Score: 23; DB: 11; Length: 29;
Best Local Similarity: 61.1%; Pred. No. 3,4e+02;
Matches: 4; Conservative: 1; Mismatches: 1; Indels: 0; Gaps: 0;

27 2 EAGRM 7
ID 11111
AC 11 EAGRM 14
PRELIMINARY: PRE: 15 AA.
D1 01 MAY 2000 (EMBLrel. 13, Created)
D1 01 MAY 2000 (EMBLrel. 14, Last sequence update)
D1 01 JUN 2000 (EMBLrel. 14, Last annotation update)
D1 17 KIA MAJOR IMMUNOPROTEIN (FRAGMENT).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archaeopteryx; Aves; Neofalcones; Galliformes; Phasianidae; Gallinae.
OX NCBI_TaxID: 9041.
RN 111
SEQUENCE:
RX MEDLINE 94022550; PubMed 7504525;
RA Yew A.W., Pearson L.R., Gaudin J.W., Bellar-Klein F.L., Bellar H.R.

```

```

R1 "An active efflux-binding domain of 17,000 daltons is associated
R1 following limited proteolysis of chicken thymus isoprenyl
R1 Biochemistr. 32:12571-12576(1994).
R1 SEQUENCE: 15 AA: 1606 BW: 6269742/98011071 CR64.

Query Match: 61.1% Score: 22; DB: 15; Length: 15;
Best Local Similarity: 80.0%; Pred. No. 2,3e+02;
Matches: 4; Conservative: 0; Mismatches: 1; Indels: 0; Gaps: 0;

QY 1 EAGM 5
ID 1111
AC 2 EAGM 6
PRELIMINARY: PRE: 20 AA.
D1 01 MAY 2000 (EMBLrel. 13, Created)
D1 01 MAY 2000 (EMBLrel. 15, Last sequence update)
D1 01 OCT 2000 (EMBLrel. 15, Last annotation update)
DE CYTOCHROME C OXIDASE SUBUNIT VIA 1 (FRAGMENT).
OS ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Bovidae;
OC Bovidae; Caprinae; ovis.
OX NCBI_TaxID: 9940.
RN 111
SEQUENCE:
RX MEDLINE 96092035; PubMed 8529022;
RA Lindor D., Freund R., Kadenbach B.
R1 "Specific expression of cytochrome c oxidase isoenzymes."
R1 Comp. Biochem. Physiol. 112B:461-469(1995).
R1 NON_TER 1
R1 NON_TER 20
R1 SEQUENCE 20 AA: 2079 BW: 5719339/6327626 CR64.

Query Match: 61.1% Score: 22; DB: 6; Length: 20;
Best Local Similarity: 66.7%; Pred. No. 3,4e+02;
Matches: 4; Conservative: 1; Mismatches: 1; Indels: 0; Gaps: 0;

QY 2 EAGRM 7
ID 1111
AC 8 EAGRM 14
PRELIMINARY: PRE: 20 AA.
D1 01 MAY 2000 (EMBLrel. 13, Created)
D1 01 MAY 2000 (EMBLrel. 14, Last sequence update)
D1 01 OCT 2000 (EMBLrel. 15, Last annotation update)
DE CYTOCHROME C OXIDASE SUBUNIT VIA 1 (FRAGMENT).
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID: 9615.
RN 111
SEQUENCE:
RX MEDLINE 96092035; PubMed 8529022;
RA Lindor D., Freund R., Kadenbach B.
R1 "Specific expression of cytochrome c oxidase isoenzymes."
R1 Comp. Biochem. Physiol. 112B:461-469(1995).
R1 NON_TER 1
R1 NON_TER 20
R1 SEQUENCE 20 AA: 2109 BW: 9219371/9752726 CR64.

Query Match: 61.1% Score: 22; DB: 6; Length: 20;
Best Local Similarity: 66.7%; Pred. No. 3,4e+02;

```




RESULT 4
 PQ0143
 C:Species: tetrahiera ornatus (evening primrose)
 C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 06 Dec 1996
 C:Accession: PQ0143
 R:Brown, S.M., Crouch, M.L.
 Plant Cell 2: 263-274, 1990
 A:Title: Characterization of a gene family abundantly expressed in *Tetrahiera ornatus*
 A:Reference number: J00992; M010:94005658
 A:Accession: PQ0143
 A:Molecule type: mRNA
 A:Residues: 1-22 (BBR)
 A:Superfamily: pollen
 C:Comment: This protein is specifically translated in the pollen.
 C:Keywords: This protein functions by depositing pollen in the pollen wall of the plant
 C:Keywords: glycosidase; hydrolase; polysaccharide degradation
 Query Match 27.2% Score 29; Db 2; Length 22;
 Best Local Similarity 93.3% Pred. No. 2.8e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 4 GAGAGG 4
 Lb 15 GAKGAG 20
 RESULT 4
 A55244
 C:Species: Paramesium tetraurelia (fragment)
 C:Date: 27-Jan-1995 #sequence_revision 27-Jan-1995 #text_change 07-Dec-1999
 C:Accession: A55244
 R:Kedeker, V., Le Gall, J.M., Schmitter, J.M., Le Gall, J.P., Fossier, J., Le Gall, J.
 Science 266: 1688-1691, 1994
 A:Title: Polyglycylation of tubulin: a posttranslational modification in axonemal microtubules
 A:Reference number: A55244; M010:95084156
 A:Accession: A55244
 A:Molecule type: protein
 A:Status: preliminary
 C:Comment: This protein is a small, water-soluble apolipoprotein, is thought to increase a residual pathway for flight-related lipid transport.
 C:Keywords: hemolymph; lipid transport
 A:Genetic code: SGR5
 C:Superfamily: tubulin
 C:Keywords: microtubule
 Query Match 25.2% Score 27; Db 2; Length 26;
 Best Local Similarity 60.9% Pred. No. 6.9e+02;
 Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
 QY 4 PAAAGGAGGAGGA 14
 Lb 15 PAAAGGAGGAGGA 26
 RESULT 5
 A55244
 C:Species: Paramesium tetraurelia (fragment)
 C:Date: 05-Jun-1997 #sequence_revision 05-Jun-1997 #text_change 25 Oct 1996
 C:Accession: A55244; A55244; A55244; B55244
 R:Kedeker, V., Le Gall, J.M., Schmitter, J.M., Le Gall, J.P., Fossier, J., Le Gall, J.
 Thromb. Res. 7: 599-610, 1975
 A:Title: Studies of the structure of canine fibrinogen.
 A:Reference number: A55244; M010:7630179
 A:Accession: A55244
 A:Molecule type: protein

A:Residues: 1-28 (BBR)
 R:Blomback, B.; Blomback, M.; Girondehl, N.J.
 Acta Chem. Scand. 19: 1789-1791, 1965
 A:Title: Studies on fibrinopeptides from mammals.
 A:Reference number: A03118
 A:Accession: A03118
 A:Molecule type: protein
 A:Residues: 1-16 (BBR)
 R:Osbaugh Jr., A.J.; Colman, R.W.; Laki, K.; Gladner, J.A.
 Biochem. Biophys. Res. Commun. 14: 555-558, 1964
 A:Reference number: A37511; M010:66020594
 A:Accession: A37511
 A:Molecule type: protein
 A:Residues: 1-16 (BBR)
 C:Superfamily: fibrinogen alpha chain; fibrinogen disulfide ring homology
 C:Keywords: blood coagulation; liver; phosphoprotein; plasma
 F:1-16 (BBR); fibrinopeptide A #status: experimental (APF)
 F:1-16 (BBR); fibrinopeptide A (partially) #status: experimental (APF)
 Query Match 25.2% Score 27; Db 2; Length 28;
 Best Local Similarity 62.5% Pred. No. 7.4e+02;
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 5 AAGAGGAGG 12
 Lb 10 AAGAGGAGG 17
 RESULT 6
 A60295
 C:Species: Acheta domestica (house cricket)
 C:Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 18-Jun-1994
 C:Accession: A60295
 R:Kedeker, V., Le Gall, J.M., Schmitter, J.M., Le Gall, J.P., Fossier, J., Le Gall, J.
 Insect Biochem. 20: 859-863, 1990
 A:Title: Adipokinetic hormone: formation of a low-density lipophorin in the hemolymph
 A:Reference number: A60295
 A:Accession: A60295
 A:Molecule type: protein
 A:Residues: 1-29 (BBR)
 C:Comment: This protein, a small, water-soluble apolipoprotein, is thought to increase a residual pathway for flight-related lipid transport.
 C:Keywords: hemolymph; lipid transport
 Query Match 24.3% Score 26; Db 2; Length 20;
 Best Local Similarity 45.5% Pred. No. 7.5e+02;
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 9 GAGGAGGAGG 19
 Lb 3 GAGGAGGAGG 13
 RESULT 7
 A55244
 C:Species: Paramesium tetraurelia (fragment)
 C:Date: 27-Jan-1995 #sequence_revision 27-Jan-1995 #text_change 07-Dec-1999
 C:Accession: A55244
 R:Kedeker, V., Le Gall, J.M., Schmitter, J.M., Le Gall, J.P., Fossier, J., Le Gall, J.
 Science 266: 1688-1691, 1994
 A:Title: Polyglycylation of tubulin: a posttranslational modification in axonemal microtubules
 A:Reference number: A55244; M010:95084156
 A:Accession: A55244
 A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 1-26 (BBR)
 C:Keywords: hemolymph; lipid transport
 A:Genetic code: SGR5
 C:Superfamily: tubulin

RESULT 14

S43045

Translation elongation factor EF-1 alpha - rabbit

C:Species: cyrtolomus carolinus (domestic rabbit)

C:Date: 17-Mar-1997 #sequence_revision 19 Mar 1997 #text_change 07 May 1999

C:Accession: S43045

R:Residuary: 144; Krall, J.A.; Devor, T.E.; Haas, R.; Louvard, D.; Morrick, W.C.

J. Biol. Chem. 264: 7096-7099, 1989

A:Title: Purification and incorporation of (3)H]ethionine into protein synthesis elongation

A:Reference number: A33048; MUID:89214136

A:Accession: S43045

A:Status: preliminary

A:Molecule type: protein

A:Residues: 119 spots

Query Match

Best Local Similarity 23.8%; Score 25.5; DB 2; Length 19;

Matches 7; Conservative 0; Mismatches 4; Indels 3; Gaps 1;

27 7 GAGGKVAIVNMF 19

17 1111

14 7 GP NWGFNVKF 16

RESULT 15

P10725

T-cell receptor beta chain V-beta region (140-21) - mouse (fragment)

C:Species: Mus musculus (house mouse)

C:Date: 17-Jul-1992 #sequence_revision 17 Jul 1992 #text_change 30-May-1997

C:Accession: P10725

R:Residuary: A.1

J. Exp. Med. 174: 115-124, 1991

A:Title: Junctional sequences of total T cell receptor beta chains have few N regions.

A:Reference number: P10509; MUID:91277601

A:Accession: P10725

A:Status: translation not shown

A:Molecule type: RNA

A:Residues: 18 -PER-

A:Experimental source: newborn thymus, strain BALB/c

A:Keywords: T-cell receptor

Query Match

Best Local Similarity 23.4%; Score 25; DB 2; Length 8;

Matches 5; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

27 5 AGGAGGK 12

17 1111

14 1 ASGAGGG 8

RESULT 16

P10725

T-cell receptor beta chain (P15) - mouse (fragment)

C:Species: Mus musculus (house mouse)

C:Date: 17-Jul-1992 #sequence_revision 17 Jul 1992 #text_change 05 May 1999

C:Accession: P10725

R:Residuary: A.1

J. Exp. Med. 174: 1471-1483, 1991

A:Title: T-cell receptor genes in a series of class I major histocompatibility complex

allelic exclusion and antigen-specific repertoire.

A:Reference number: P10725; MUID:92078846

A:Accession: P10725

A:Molecule type: mRNA

A:Residues: 1-16 -CASS-

A:Experimental source: EMBL X65857, NIH 3T3/3247.1, F10-3247.4

A:Keywords: T-cell receptor

Best Local Similarity 66.7%; Pred. No. 8.5e+02;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PPRGAQ 6

14 1111

14 5 PRQGNQ 10

Search completed: May 1, 2001, 11:26:21

Job time: 126 sec

Query Match

Best Local Similarity 23.4%; Score 25; DB 2; Length 16;

D1 01 FEB 1991 (Rel. 17, Last sequence update)
 D1 01 FEB 1991 (Rel. 17, Last annotation update)
 DE LIPOLASE INHIBITOR (FRAGMENT).
 CC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 CC Insecta; Neoptera; Orthoptera; Dictyoptera; Blattaria;
 CC Blattaria; Mantodea; Dermaptera; Dermaptera;
 CC NCBI TaxID 6176;
 RN [1]
 RP SEQUENCE.
 AC 11350E Homologous
 PX MEDLINE 8455217; Pubmed 2777794.
 RA Baskova Y., Ishikawa A.P., Yepich M., Deemer R.G.H.;
 R1 "Participation and characterization of lipolase inhibitor from
 R1 hemolymph of the American cockroach."
 R1 J. Biol. Chem. 264:16165-16169 (1989).
 CC 1. FUNCTION: UNDER RESTRICTED CIRCUMSTANCES, INHIBITS LIPOLASE IN A
 CC DOSE DEPENDENT MANNER.
 DE EMBL A54233; A4233.
 KW Glycoprotein.
 FI NON TER.
 SQ SEQUENCE 29 AA: 2934 MW: 4574737AFVGA22 "R%4.

Query Match 21.9% Score 23; DB 1; Length 19;
 Best local similarity 6.2% Pred. No. 1.7e-04;
 Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0.

27 5.064668
 111 1
 13 6.064668 19

RESULT 12
 ML2 BRANA
 AC 596653
 D1 01 NOV 1997 (Rel. 55, Created)
 D1 01 NOV 1997 (Rel. 55, Last sequence update)
 D1 01 NOV 1997 (Rel. 55, Last annotation update)
 DE METALLOTHIONEIN (HEAVY CHAIN TYPE 2, 23-215) (FRAGMENT).
 RN 152410.
 CC Eukaryota; Eukaryota; Embryophyta; Tracheophyta; Spermatophyta;
 CC Eukaryota; Vascular plantae; Embryophyta; Tracheophyta; Spermatophyta;
 CC Monocotyledonae; Eudicotyledones; Core eudicots; Rosidae; eurosids 1;
 CC Brassicales; Brassicales; Brassicae;
 CC NCBI TaxID 676;
 RN [1]
 RP SEQUENCE FROM N.A.
 AC STRAIN CV, FALCON; TISSUE Leaf.
 RA Buchanan-Wallston, V., Ainsworth C.,
 CC submitted (Sep 1999) to the EMBL/GenBank/DDBJ databases.
 CC 1. FUNCTION: METALLOTHIONEINS HAVE A HIGH CONTENT OF CYSTEINE
 CC RESIDUES THAT BIND VARIOUS HEAVY METALS.
 CC 1. SIMILARITY: BELONGS TO FAMILY 15 IN METALLOTHIONEIN SUPERFAMILY.

This SWISS prot entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see <http://www.isb.slb.ch/announcements>
 or send an email to license@isb.slb.ch).

EMBL 066222; AAC5104.1;
 KW Metal binding; Metal thiolate cluster; Chelation.
 FI NON TER.
 SQ SEQUENCE 28 AA: 2913 MW: 94696AUB414F04A "R%64.

Query Match 21.9% Score 23; DB 1; Length 28;
 Best local similarity 42.9% Pred. No. 1.6e-04;
 Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0.

QY 5 AGGAGCR 11
 111 1
 DE 9 AENDACK 15
 RESULT 13
 MDL BURS
 D1 MDL BURS STANDARD: PRI: 29 AA.
 AC 980536;
 D1 01-FEB-1996 (Rel. 43, Created)
 D1 01 FEB 1996 (Rel. 43, Last sequence update)
 D1 01 OCT 2000 (Rel. 40, Last annotation update)
 DE MALATE DEHYDROGENASE (P-1.1.1.37) (FRAGMENT).
 RN MDL.
 CC Burkholderia pseudomallei (Pseudomonas pseudomallei).
 CC Bacteriia; Proteobacteria; Beta superdivisiob; Burkholderia group;
 CC Burkholderia.
 CC NCBI TaxID 26450;
 RN [1]
 RP SEQUENCE.
 AC STRAIN A133 2343;
 RX MEDLINE 9734112; Pubmed 9190829;
 RA Charnock C.;
 R1 "Structural studies of malate dehydrogenases (MDHs): MDHs in
 R1 Brownian motion species are the first reported MDHs in Proteobacteria
 R1 which resemble lactate dehydrogenases in primary structure."
 R1 J. Bacteriol. 179:4066-4070 (1997).
 CC 1. CATALYTIC ACTIVITY: L-MALATE + NAD(+) -> OXALOACETATE + NADH.
 CC 1. SIMILARITY: BELONGS TO THE MDH FAMILY, MDH SUBFAMILY.
 DE HSSP: P11708; 4MDH.
 DR InterPro: IPR001246;
 DR InterPro: IPR001252;
 DR Pfam: PF00056; MDH; 1.
 DR PROSITE: PS00066; MDH; PARTIAL.
 KW Glutamate dehydrogenase, Tricarballic acid cycle, NAD.
 FI NON TER.
 SQ SEQUENCE 29 AA: 3046 MW: 1042674E7C7AA94 "R%64.

Query Match 21.9% Score 23; DB 1; Length 29;
 Best local similarity 31.2% Pred. No. 1.7e-04;
 Matches 5; Conservative 5; Mismatches 2; Indels 4; Gaps 1.

QY 4 GAGGACGCAVAVNPV 19
 111 1
 DE 11 GAGGQ----TAYSLLE 22

RESULT 14
 ML2 AVAB
 D1 ML2 AVAB STANDARD: PRI: 27 AA.
 AC 904548;
 D1 20-MAR-1987 (Rel. 04, Created)
 D1 20-MAR-1987 (Rel. 04, Last sequence update)
 D1 01-NOV-1997 (Rel. 45, Last annotation update)
 DE METALLOTHIONEIN (MD).
 CC Eukaryota; Eukaryota; Embryophyta; Tracheophyta; Spermatophyta;
 CC Eukaryota; Vascular plantae; Embryophyta; Tracheophyta; Spermatophyta;
 CC Monocotyledonae; Eudicotyledones; Core eudicots; Rosidae; eurosids 1;
 CC Brassicales; Brassicales; Brassicae;
 CC NCBI TaxID 5441;
 RN [1]
 RP SEQUENCE.
 AC STRAIN A 42; TISSUE Mycelium;
 RA Monier K., Lerch K.;
 R1 "Copper metallothionein from the fungus Arctium bisporus: chemical
 R1 and spectroscopic properties."
 R1 Biochemistry 24:6751-6756 (1985).
 CC 1. MISCELLANEOUS: THE SEVEN CYSTEINES BIND SIX COPPER (CHROMIUM) IONS.
 CC 1. MISCELLANEOUS: THE SEVEN CYSTEINES BIND SIX COPPER (CHROMIUM) IONS.
 CC PROPERTIES OF THE COPPER METAL ARE ATTRIBUTED TO THE METAL THIOLE
 CC COMPLEX BECAUSE THEY ARE NOT PRESENT IN THE APOPROTEIN.
 CC 1. SIMILARITY: BELONGS TO FAMILY 8 IN METALLOTHIONEIN SUPERFAMILY.

DB PIP: A02286; SMMR.
 KW Metal-binding: Metal-thiolate cluster; Chelation: Copper.
 SC SEQUENCE: 25 AA; 224 MW; CAP20121020P7 CPO64;
 Query Match: 21.0%; Score 22.5; DB 1; Length 25;
 Best Local Similarity 55.6%; Pred. No. 1.8e+03;
 Matches 5; Conservative 1; Mismatches 2; Indels 1; Gaps 1;
 QY 7 GD-GRGVA 14
 L 1 1 1 1
 LR 1 GDCGSCAS 9
 RESULTS 15
 MI_NMRCK
 ID M. MERRY STACWAPL PRT 25 AA.
 AC P22697;
 DI 21-JUL-1996 (Ref., 01; Created)
 DI 21-JUL-1986 (Ref., 01; Last sequence update)
 DI 15-JUL-1999 (Ref., 08; Last annotation update)
 DE METALLO-THIONEIN (MT).
 GN MT.
 OS Neurospora crassa.
 SC Escherichia: Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 SC Sordariales; Sordariaceae; Neurospora.
 X N94_14310-5141;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE 96040247; PubMed 2942441;
 RA Mueller K., Germain D.A., Lerch K.;
 RI "Isolation and structural organization of the Neurospora crassa
 RI copper metallothionein gene."
 R1 EMBO J. 3: 4:266-268(1985).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE 91112114; PubMed 2148862;
 RA Beyer F.M., Jacobs F.A., Brausseau P.;
 RI "Expression of a Neurospora crassa metallothionein and its variants
 RI in Escherichia coli."
 R1 A4 J. Environ. Microbiol. 56:2748-2754(1990).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE 92029909; PubMed 2959528;
 RA Mueller K., Germain D.A., Lerch K.;
 RI "Isolation and regulation of expression of the Neurospora crassa
 RI copper metallothionein gene."
 R1 Experimental Suppl. 52:934-100(1997).
 RN [4]
 RP SEQUENCE.
 RX MEDLINE 9014244; PubMed 6444697;
 RA Lerch K.;
 RI "Copper metallothionein, a copper-binding protein from Neurospora
 RI crassa."
 R1 Nature 284: 368-370(1980).
 RN [5]
 RP SEQUENCE BY NMR.
 RX MEDLINE 89382847; PubMed 2525920;
 RA Malikyan J.A., Lerch K., Amilage J.M.;
 RI "Protein NMR studies of a metallothionein from Neurospora crassa:
 RI sequence-specific assignments by ROE measurements in the rotating
 RI frame."
 R1 Biochemistry 28:2941-2955(1989).
 CC -1 INDUCTION: BY COPPER.
 CC -1 MISCELLANEOUS: THE GIVEN CYSTEINES BIND SIX COPPER (SUPEROXIDE) IONS.
 CC -1 SIMILARITY: BELONGS TO FAMILY 8 IN METALLOTHIONEIN SUPERFAMILY.
 CC This SWISS-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license from the Swiss Institute of Bioinformatics.

CC or send an email to license@sib.ch).
 CC ---
 DB EMBL: X03009; CAA26794.1;
 DB EMBL: M59836; AAA48594.1;
 DB EMBL: M27709; AAA43595.1;
 DB PIR: A24641; SMMR.
 KW Metal-binding: Metal thiolate cluster; Chelation: Copper.
 FT INIT_MET 0 0
 FT METAL 3 4 COPPER.
 FT METAL 5 5
 FT METAL 11 11 COPPER.
 FT METAL 14 14 COPPER.
 FT METAL 17 17 COPPER.
 FT METAL 19 19 COPPER.
 FT METAL 22 22 COPPER.
 FT CONFLICT 21 21 N 1 (IN REF. 2).
 CC MISCHE 25 AA; 224 MW; 256 RC201022P7 CPO64;
 Query Match: 21.0%; Score 22.5; DB 1; Length 25;
 Best Local Similarity 55.6%; Pred. No. 1.8e+03;
 Matches 5; Conservative 1; Mismatches 2; Indels 1; Gaps 1;
 QY 7 GD-GRGVA 14
 L 1 1 1 1
 DB 1 GDCGSCAS 9

Search completed: May 1, 2001, 11:25:54
 Job time: 101 sec



Genome version 4.5
Copyright (c) 1994-2000 Compugen Ltd.

OM protein protein search using SW model

Run on: May 1, 2001, 11:27:05, Search time 40.82 seconds
(without alignments)
45,941 Million cell updates/sec

File: us-09-341-829a_5_copy_164_179
Port: 82
Sequence: 1 KARDITIKIRVSEQR 16

Scoring table: 45 SUM62
Gap: 16.0, Gapex: 0.1

Search: 3400 seqs, 11720715 residues

Total number of hits satisfying chosen parameters: 10868

Minimum db seq length: 0
Maximum db seq length: 50

Post processing: Maximum Match 99
Minimum Match 100%
Cutoff 1181 4% summaries

Database:

- SPIRITML_15: *
- 1: sp. archaea: *
- 2: sp. bacteria: *
- 3: sp. fungi: *
- 4: sp. human: *
- 5: sp. invertebrates: *
- 6: sp. mammals: *
- 7: sp. birds: *
- 8: sp. ornithines: *
- 9: sp. phages: *
- 10: sp. plant: *
- 11: sp. eukarya: *
- 12: sp. metazoa: *
- 13: sp. vertebrates: *
- 14: sp. virus: *

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed.
and is derived by analysis of the total score distribution

SUMMARIES

Prod. No.	Score	Match	Length	db	ID	Description
1	26	41.7	21	2	Q9X117	Q9X117, Prochlorococcus
2	26	41.7	21	14	Q92084	Q92084, human rhino
3	26	41.7	27	6	Q91626	Q91626, sus. serofa
4	26	41.7	26	4	Q91626	Q91626, homo sapien
5	26	41.7	25	7	Q91626	Q91626, homo sapien
6	26	41.7	27	14	Q91626	Q91626, homo sapien
7	26	41.7	27	14	Q91626	Q91626, homo sapien
8	26	41.7	27	14	Q91626	Q91626, homo sapien
9	26	41.7	27	14	Q91626	Q91626, homo sapien
10	26	41.7	27	14	Q91626	Q91626, homo sapien
11	26	41.7	27	14	Q91626	Q91626, homo sapien
12	26	41.7	27	14	Q91626	Q91626, homo sapien
13	26	41.7	27	14	Q91626	Q91626, homo sapien
14	26	41.7	27	14	Q91626	Q91626, homo sapien
15	26	41.7	27	14	Q91626	Q91626, homo sapien
16	26	41.7	27	14	Q91626	Q91626, homo sapien
17	26	41.7	27	14	Q91626	Q91626, homo sapien
18	26	41.7	27	14	Q91626	Q91626, homo sapien
19	26	41.7	27	14	Q91626	Q91626, homo sapien

ALIGNMENTS

RESULT 1

Q9X117	1	26	41.7	27	14	Q9X117
10	Q9X117					
AC	Q9X117					
DT	01-NOV-1999 (TEMBLREL, 12, Created)					
DT	01-NOV-1999 (TEMBLREL, 12, Last sequence update)					
DT	01-NOV-1999 (TEMBLREL, 12, Last annotation update)					
DE	CYTCHROME B6/F COMPLEX SUBUNIT IV (FRAGMENT)					
GN	PEPD					
OS	Prochlorococcus sp.					
OC	Bacteria; Cyanobacteria; Prochlorophytes; Prochlorococcus					
OX	NCBI TaxID: 1220					
FN	11					
RP	SEQUENCE FROM N.A.					
RA	Urbach E., Chisholm S.W.					
RT	Sorted by diversity in Prochlorococcus populations from the Sargasso Sea and Gulf Stream.					
RL	Limol., oceanic, 48:1615-1640(1998)					
DR	EMBL: AF070208; A024258.1					
FI	NON-REF					
SI	SEQUENCE 21 AA; 2460 MW; Q9X117:27-27F Q9X117					
Query Match						
Best Local Similarity 50.00%; Score 40; DB 2; Length 21						
Matches R: Consistency 0; Mismatches R: Indels 0; Gaps 0						
QY	1 KARDITIKIRVSEQR 16					
DB	5 KARDITIKIRVSEQR 20					
RESULT 2						
Q92084						
10	Q92084					
AC	Q92084					
DT	01-NOV-1996 (TEMBLREL, 01, Created)					
DT	01-NOV-1996 (TEMBLREL, 01, Last sequence update)					
DT	01-NOV-1998 (TEMBLREL, 08, Last annotation update)					
DE	TYPE 1A (HRV 1A) GENOMIC LINKED VIRAL PROTEIN (VING) GENOMIC REGION (FRAGMENT)					

```

OS Homo rhinovirus
CA Virus: ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Rhinovirus.
OX NCBI_TaxID: 12133;
RN 1;
RA Sequence from N.A.
RE MEDLINE: 95144022; PubMed 3512851;
KA Watson J., Watson J.B., Raney E., Saffert J.M., Werner F.J.,
  Resmer J.;
PI "Molecular cloning and sequence determination of the genomic regions
  encoding protease and genome-linked protein of three picornaviruses."
  J. Virol. 57:1084-1093(1986).
DK EMBL: M12166; AAA45760.1;
FI NON-REF 1;
FI R-REF 21;
SQ SEQUENCE 21 AA: 2424 MW: 500E171ED80942 CRC64;

Query Match 35.4%; Score 29; DB 14; Length 21;
Best Local Similarity 66.7%; Pred. No. 1.8e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 6 PRKRVSEQR 16
  1 1 1 1 1
  6 PRKRVSEQR 17

RESULT 5
Q91RZ6 PRELIMINARY: PRT: 27 AA.
AC Q91RZ6;
DI 01-MAY-2000 (TrEMBLrel. 13, Created)
DI 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DI 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE ADRENOMEDULLIN.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.
OX NCBI_TaxID: 9824;
RN 1;
RA Sequence from N.A.
RE MEDLINE: 95157714; PubMed 8576091;
RA Ehlert Y., Fildes F., Fildes F., Sawamoto M., Matsuo H., Eto T.;
  "Distribution and characterization of immunoreactive adrenomedullin in
  porcine tissue, and isolation of adrenomedullin [26-52] and
  adrenomedullin [34-52] from porcine duodenum."
  J. Biochem. 118:765-770(1995).
DK INTERFERON PRO001710;
DE PEPTIDES: PRO0001; ADRENOMEDULLIN.
SQ SEQUENCE 27 AA: 4363 MW: 8B8A7FA18DBB4050 CRC64;

```

```

Query Match 24.1%; Score 28; DB 5; Length 27;
Best Local Similarity 40.0%; Pred. No. 3.5e+02;
Matches 6; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 1 KAPNPKPKVSEQR 15
  1 1 1 1 1 1
  11 KAPNPKPKVSEQR 25

RESULT 4
Q9NFX0 PRELIMINARY: PRT: 28 AA.
AC Q9NFX0;
DI 01-OCT-2000 (TrEMBLrel. 15, Created)
DI 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DI 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE BOLL118M15.1.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

```

```

OX NCBI_TaxID: 9606;
RN 1;
RA Sequence from N.A.
RE MEDLINE: 95144022; PubMed 3512851;
KA Watson J., Watson J.B., Raney E., Saffert J.M., Werner F.J.,
  Resmer J.;
PI "Molecular cloning and sequence determination of the genomic regions
  encoding protease and genome-linked protein of three picornaviruses."
  J. Virol. 57:1084-1093(1986).
DK EMBL: M12166; AAA45760.1;
FI NON-REF 1;
FI R-REF 21;
SQ SEQUENCE 28 AA: 2226 MW: 40EAF2490B199 CRC64;

Query Match 44.1%; Score 28; DB 4; Length 28;
Best Local Similarity 38.5%; Pred. No. 3.6e+02;
Matches 5; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 4 LKPKPKVSEQR 16
  1 1 1 1 1
  7 DGRSPEREKQK 19

RESULT 5
Q9TNQ6 PRELIMINARY: PRT: 25 AA.
AC Q9TNQ6;
DI 01-MAY-2000 (TrEMBLrel. 13, Created)
DI 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DI 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE HLA CLASS II-ASSOCIATED INVARIANT CHAIN 11.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID: 9606;
RN 1;
RA Sequence.
RE MEDLINE: 92375195; PubMed 1389674;
RA Chicz R.M., Urban R.G., Lane W.S., Gorra J.C., Stern L.J.,
  Vignali D.A., Strominger J.L.;
  "Predominant naturally processed peptides bound to HLA DR1 are derived
  from MHC-related molecules and are heterogeneous in size."
  Nature 358:764-768(1992).
DK EMBL: 358764 768(1992);
KW MHC.
SQ SEQUENCE 25 AA: 2793 MW: 2F80189276A802F4 CRC64;

Query Match 31.7%; Score 26; DB 7; Length 25;
Best Local Similarity 50.0%; Pred. No. 7.3e+02;
Matches 6; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 5 LKPKPKVSEQR 16
  1 1 1 1 1
  1 LKPKPKVSEKMR 12

RESULT 6
Q9Q170 PRELIMINARY: PRT: 27 AA.
AC Q9Q170;
DI 01-MAY-2000 (TrEMBLrel. 13, Created)
DI 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DI 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE POLYPROTEIN (FRAGMENT).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID: 11103;
RN 1;
RA Sequence from N.A.
RE Sandres K., Dubois M., Pasquier C., Izopet J.;
  "The genetic heterogeneity of hypervariable region 1 of the viral
  RNA and the sensitivity of hepatitis C virus to interferon alpha
  therapy."
  J. Clin. Virol. 1995;
DK EMBL: AF165797; AA052365.1;
DE INTERPRO: IPR002531;

```

DE FRAM: PF01560; HCV_NSI: 1;
 KW Polyprotein.
 FT NON_TER 1 1
 FT NON_TER 27 27
 SQ SEQUENCE 27 AA: 2945 MW: 91184189A752H130 CR664;

Query Match: 41.7%; Score 26; 108 14; Length 27;
 Best Local Similarity: 83.9%;
 Matches: 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 ARDLRT 7
 11111
 DB 10 ARDVRT 15

RESULT 7
 Q90167 PRELIMINARY: PRT: 27 AA.

AC Q90167; 01 MAY 2000 (TrEMBLrel. 14, Created)
 DI 01 MAY 2000 (TrEMBLrel. 14, Last sequence update)
 DI 01 JUN 2000 (TrEMBLrel. 14, Last annotation update)
 DE POLYPROTEIN (FRAGMENT).
 OS Hepatitis C virus.
 CC Viruses; ssRNA positive-strand viruses, no RNA stage; Flaviviridae;
 CC Hepacivirus.
 GX NCBI_TaxID:11103;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Sandrois K., Iudais M., Pasquier C., Lopez J.,
 RT "The genetic heterogeneity of hypervariable region 1 of the viral
 genome and the sensitivity of hepatitis C virus to interferon alpha
 therapy.";
 RL Submitted (2001-1999) to the EMBL/GenBank/TrEMBL databases.
 DE EMBL: AF166710; AA025468.1;
 DE INTERPRO: IPR002531;
 DE FRAM: PF01560; HCV_NSI: 1;
 KW Polyprotein.
 FT NON_TER 1 1
 FT NON_TER 27 27
 SQ SEQUENCE 27 AA: 2945 MW: 91184189A752H130 CR664;

Query Match: 41.7%; Score 26; 108 14; Length 27;
 Best Local Similarity: 83.9%;
 Matches: 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 ARDLRT 7
 11111
 DB 10 ARDVRT 15

RESULT 8
 Q90167 PRELIMINARY: PRT: 27 AA.

AC Q90167; 01 MAY 2000 (TrEMBLrel. 14, Created)
 DI 01 MAY 2000 (TrEMBLrel. 14, Last sequence update)
 DI 01 JUN 2000 (TrEMBLrel. 14, Last annotation update)
 DE POLYPROTEIN (FRAGMENT).
 OS Hepatitis C virus.
 CC Viruses; ssRNA positive-strand viruses, no RNA stage; Flaviviridae;
 CC Hepacivirus.
 GX NCBI_TaxID:11103;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Sandrois K., Iudais M., Pasquier C., Lopez J.,
 RT "The genetic heterogeneity of hypervariable region 1 of the viral
 genome and the sensitivity of hepatitis C virus to interferon alpha
 therapy.";
 RL Submitted (2001-1999) to the EMBL/GenBank/TrEMBL databases.
 DE EMBL: AF166710; AA025468.1;

DE FRAM: PF01560; HCV_NSI: 1;
 KW Polyprotein.
 FT NON_TER 1 1
 FT NON_TER 27 27
 SQ SEQUENCE 27 AA: 2945 MW: 91184189A752H130 CR664;

Query Match: 41.7%; Score 26; 108 14; Length 27;
 Best Local Similarity: 83.9%;
 Matches: 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 ARDLRT 7
 11111
 DB 10 ARDVRT 15

RESULT 9
 Q90167 PRELIMINARY: PRT: 27 AA.

AC Q90167; 01 MAY 2000 (TrEMBLrel. 14, Created)
 DI 01 MAY 2000 (TrEMBLrel. 14, Last sequence update)
 DI 01 JUN 2000 (TrEMBLrel. 14, Last annotation update)
 DE POLYPROTEIN (FRAGMENT).
 OS Hepatitis C virus.
 CC Viruses; ssRNA positive-strand viruses, no RNA stage; Flaviviridae;
 CC Hepacivirus.
 GX NCBI_TaxID:11103;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Sandrois K., Iudais M., Pasquier C., Lopez J.,
 RT "The genetic heterogeneity of hypervariable region 1 of the viral
 genome and the sensitivity of hepatitis C virus to interferon alpha
 therapy.";
 RL Submitted (2001-1999) to the EMBL/GenBank/TrEMBL databases.
 DE EMBL: AF166710; AA025468.1;
 DE INTERPRO: IPR002531;
 DE FRAM: PF01560; HCV_NSI: 1;
 KW Polyprotein.
 FT NON_TER 1 1
 FT NON_TER 27 27
 SQ SEQUENCE 27 AA: 2945 MW: 91184189A752H130 CR664;

Query Match: 41.7%; Score 26; 108 14; Length 27;
 Best Local Similarity: 83.9%;
 Matches: 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 ARDLRT 7
 11111
 DB 10 ARDVRT 15

RESULT 10
 Q90167 PRELIMINARY: PRT: 27 AA.

AC Q90167; 01 MAY 2000 (TrEMBLrel. 14, Created)
 DI 01 MAY 2000 (TrEMBLrel. 14, Last sequence update)
 DI 01 JUN 2000 (TrEMBLrel. 14, Last annotation update)
 DE POLYPROTEIN (FRAGMENT).
 OS Hepatitis C virus.
 CC Viruses; ssRNA positive-strand viruses, no RNA stage; Flaviviridae;
 CC Hepacivirus.
 GX NCBI_TaxID:11103;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Sandrois K., Iudais M., Pasquier C., Lopez J.,
 RT "The genetic heterogeneity of hypervariable region 1 of the viral
 genome and the sensitivity of hepatitis C virus to interferon alpha
 therapy.";
 RL Submitted (2001-1999) to the EMBL/GenBank/TrEMBL databases.

EMBL: AF166716; AA052474.1;
 InterPro: IPR002541;
 Pfam: PF01540; HCV_NS1; 1;
 PolyPhen: 1;
 N N 119 1
 E E N N 119 27 27
 S: SEQUENCE 27 AA: 2935 MW: 91104099A7628130 CRC64;

Query Match: 41.7%; Score 26; db 14; Length 27;
 Best Local Similarity: 83.0%; Prod. No: 7.9e+02;

Matches: 5; Conservative: 1; Mismatches: 0; Indels: 0; Gaps: 0;

27 1. APLGRLT
 11 11111
 16 16 ADYVRL 15

PREMIL 1
 Query 1
 16 Q92F61 PRELIMINARY; PPT; 27 AA.
 AC Q92F61;
 DI 07 MAY 2000 (1:EMBLrel. 13, Created)
 DI 01 MAY 2000 (1:EMBLrel. 13, Last sequence update)
 DI 01 JUN 2000 (1:EMBLrel. 14, Last annotation update)
 OS POLYOMELIN (PACHMENT)
 OS Hepatitis C virus
 NC Virus; ssRNA positive-strand viruses, no RNA stage; Flaviviridae;
 NC Hepatitis C virus
 NC NCBI TaxID: 11103;
 FN 1;
 FE SEQUENCE FROM N.A.
 RA Sandrois K., Indels M., Pasquier C., Lopez J.;
 RI "The genetic heterogeneity of hypervariable region 1 of the viral
 PI genome and the sensitivity of hepatitis C virus to interferon alpha
 R1 therapy.";
 R1 Submitted (2 JUL 1999) to the EMBL/GenBank/DBJ databases.
 LR EMBL: AF166716; AA052474.1;
 LR INTERPRO: IPR002541;
 LR Pfam: PF01540; HCV_NS1; 1;
 KW Polypeptide;
 FT N N 119 1
 FT N N 119 27 27
 S: SEQUENCE 27 AA: 2935 MW: 91104099A7628130 CRC64;

Query Match: 41.7%; Score 26; db 14; Length 27;
 Best Local Similarity: 83.0%; Prod. No: 7.9e+02;

Matches: 5; Conservative: 1; Mismatches: 0; Indels: 0; Gaps: 0;

27 2 APLGRLT
 11 11111
 16 16 ADYVRL 15

Search completed: May 1, 2001, 11:27:05
 Database: 145, size



DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
 DE POLYPROTEIN (FRAGMENT)
 OS Hepatitis C virus
 OC Viruses; ssRNA positive strand viruses; no DNA stage; Flaviviridae;
 NC Hepacivirus
 NX NCBI_TaxID:11104;
 RN 111
 RP SEQUENCE FROM N.A.
 RA Sandres K., Dubois M., Pasquier C., Lopez J.
 RI "The genetic heterogeneity of hypervariable region 1 of the viral
 RT genome and the sensitivity of hepatitis C virus to interferon alpha
 RI therapy."
 RL Submitted (JUL 1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF166837; AAU52496.1;
 DR INTERPRO: IPR002531;
 DR PFAM: PF01560; HCV_NSI: 1;
 KW Polyprotein.
 FT NON_TER 1 1
 FT NON_TER 27 27
 SQ SEQUENCE 27 AA; 2757 MW; 61AF606C429DE22 chr64.

Query Match 29.9% Score 42; DB 14; Length 27;

Best Local Similarity 43.8% Pred. No. 17602;

Matches 7; Conservative 4; Mismatches 3; Indels 2; Gaps 1;

QY 4 CAQIACGCGVA FNV 17

111 111111111

DB 7 GSEAHGAGLASLNN 22

RESULT 7

Q%QHT9

AC Q%QHT9 PRELIMINARY; PRI: 27 AA.

DT 01-MAY-2000 (TrEMBLrel. 14, Created)

DI 01-MAY-2000 (TrEMBLrel. 14, Last sequence update)

DI 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)

DE POLYPROTEIN (FRAGMENT)

OS Hepatitis C virus

OC Viruses; ssRNA positive strand viruses; no DNA stage; Flaviviridae;

NC Hepacivirus

NX NCBI_TaxID:11104;

RN 111

RP SEQUENCE FROM N.A.

RA Sandres K., Dubois M., Pasquier C., Lopez J.

RI "The genetic heterogeneity of hypervariable region 1 of the viral

RT genome and the sensitivity of hepatitis C virus to interferon alpha

RI therapy."

RL Submitted (JUL 1999) to the EMBL/GenBank/DBJ databases.

DR EMBL: AF166837; AAU52496.1;

DR INTERPRO: IPR002531;

DR PFAM: PF01560; HCV_NSI: 1

KW Polyprotein.

FT NON_TER 1 1

FT NON_TER 27 27

SQ SEQUENCE 27 AA; 2757 MW; 61AF606C429DE22 chr64.

Query Match

29.9% Score 42; DB 14; Length 27;

Best Local Similarity 43.8% Pred. No. 17602;

Matches 7; Conservative 4; Mismatches 3; Indels 2; Gaps 1;

QY 4 CAQIACGCGVA FNV 17

111 111111111

DB 7 GSEAHGAGLASLNN 22

RESULT 7

Q%QHT9

AC Q%QHT9 PRELIMINARY; PRI: 27 AA.

DT 01-MAY-2000 (TrEMBLrel. 14, Created)

DT 01-MAY-2000 (TrEMBLrel. 14, Last sequence update)
 DI 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
 DE POLYPROTEIN (FRAGMENT)
 OS Hepatitis C virus
 OC Viruses; ssRNA positive strand viruses; no DNA stage; Flaviviridae;
 NC Hepacivirus
 NX NCBI_TaxID:11104;
 RN 111
 RP SEQUENCE FROM N.A.
 RA Sandres K., Dubois M., Pasquier C., Lopez J.
 RI "The genetic heterogeneity of hypervariable region 1 of the viral
 RT genome and the sensitivity of hepatitis C virus to interferon alpha
 RI therapy."
 RL Submitted (JUL 1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF166837; AAU52496.1;
 DR INTERPRO: IPR002531;
 DR PFAM: PF01560; HCV_NSI: 1;
 KW Polyprotein.
 FT NON_TER 1 1
 FT NON_TER 27 27
 SQ SEQUENCE 27 AA; 2757 MW; 61AF606C429DE22 chr64.

Query Match 29.9% Score 42; DB 14; Length 27;

Best Local Similarity 43.8% Pred. No. 17602;

Matches 7; Conservative 4; Mismatches 3; Indels 2; Gaps 1;

QY 4 CAQIACGCGVA FNV 17

111 111111111

DB 7 GSEAHGAGLASLNN 22

RESULT 9

Q%QHT9

AC Q%QHT9 PRELIMINARY; PRI: 27 AA.

DT 01-MAY-2000 (TrEMBLrel. 14, Created)

DI 01-MAY-2000 (TrEMBLrel. 14, Last sequence update)

DI 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)

DE POLYPROTEIN (FRAGMENT)

OS Hepatitis C virus

OC Viruses; ssRNA positive strand viruses; no DNA stage; Flaviviridae;

NC Hepacivirus

NX NCBI_TaxID:11104;

RN 111

RP SEQUENCE FROM N.A.

RA Sandres K., Dubois M., Pasquier C., Lopez J.

RI "The genetic heterogeneity of hypervariable region 1 of the viral

RT genome and the sensitivity of hepatitis C virus to interferon alpha

RI therapy."

RL Submitted (JUL 1999) to the EMBL/GenBank/DBJ databases.

DR EMBL: AF166837; AAU52496.1;

DR INTERPRO: IPR002531;

DR PFAM: PF01560; HCV_NSI: 1

KW Polyprotein.

FT NON_TER 1 1

FT NON_TER 27 27

SQ SEQUENCE 27 AA; 2757 MW; 61AF606C429DE22 chr64.

Query Match

29.9% Score 42; DB 14; Length 27;

Best Local Similarity 43.8% Pred. No. 17602;

Matches 7; Conservative 4; Mismatches 3; Indels 2; Gaps 1;

QY 4 CAQIACGCGVA FNV 17

111 111111111

DB 7 GSEAHGAGLASLNN 22

RESULT 10

Q%QHT9

AC Q%QHT9 PRELIMINARY; PRI: 27 AA.

DT 01-MAY-2000 (1:EMBLrel, 13, Created)
DI 01-MAY-2000 (1:EMBLrel, 13, Last sequence update)
DI 01-JUN-2000 (1:EMBLrel, 14, Last annotation update)
DE POLYPROTEIN (FRAGMENT)
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RA Sandres K., Dubois M., Pasquier C., Izopet J.,
RI "The genetic heterogeneity of hypervariable region 1 of the viral
RT genome and the sensitivity of hepatitis C virus to interferon alpha
therapy.";
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
RE EMBL: AF166839; AAF52501.1;
DR INTERPRO: IPR002541;
DR PFAM: PF01560; HCV_NSI; 1;
KW Polyprotein.
FT NON_TER 1 1
FT NON_TER 27 27
SQ SEQUENCE 27 AA: 2757 MW: 60AF66C4295BE22 CR064;

Query Match 29.9% Score 32; DB 14; Length 27;
Best Local Similarity 43.9% Pred. No. 1.7e+02;
Matches 7; Conservative 4; Mismatches 9; Indels 2; Gaps 1;

QY 4 GAQDGRGVA--FNV 17
111 1111111
DB 7 GSEARGARGLASLFNM 22

RESULT 11

Q9QHT6 PRELIMINARY; PRT; 27 AA.
DI 01-MAY-2000 (1:EMBLrel, 13, Created)
DI 01-MAY-2000 (1:EMBLrel, 14, Last sequence update)
DE POLYPROTEIN (FRAGMENT)
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RA Sandres K., Dubois M., Pasquier C., Izopet J.,
RI "The genetic heterogeneity of hypervariable region 1 of the viral
RT genome and the sensitivity of hepatitis C virus to interferon alpha
therapy.";
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
RE EMBL: AF166841; AAF52499.1;
DR INTERPRO: IPR002541;
DR PFAM: PF01560; HCV_NSI; 1;
KW Polyprotein.
FT NON_TER 1 1
FT NON_TER 27 27
SQ SEQUENCE 27 AA: 2757 MW: 60AF66C4295BE22 CR064;

Query Match 29.9% Score 32; DB 14; Length 27;
Best Local Similarity 43.9% Pred. No. 1.7e+02;
Matches 7; Conservative 4; Mismatches 9; Indels 2; Gaps 1;

QY 4 GAQDGRGVA--FNV 17
111 1111111
DB 7 GSEARGARGLASLFNM 22

RESULT 12

Q9QHT5 PRELIMINARY; PRT; 27 AA.

Q9QHT5
DI 01-MAY-2000 (1:EMBLrel, 13, Created)
DI 01-MAY-2000 (1:EMBLrel, 13, Last sequence update)
DI 01-JUN-2000 (1:EMBLrel, 14, Last annotation update)
DE POLYPROTEIN (FRAGMENT)
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RA Sandres K., Dubois M., Pasquier C., Izopet J.,
RI "The genetic heterogeneity of hypervariable region 1 of the viral
RT genome and the sensitivity of hepatitis C virus to interferon alpha
therapy.";
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
RE EMBL: AF166842; AAF52500.1;
DR INTERPRO: IPR002541;
DR PFAM: PF01560; HCV_NSI; 1;
KW Polyprotein.
FT NON_TER 1 1
FT NON_TER 27 27
SQ SEQUENCE 27 AA: 2757 MW: 60AF66C4295BE22 CR064;

Query Match 29.9% Score 32; DB 14; Length 27;
Best Local Similarity 43.9% Pred. No. 1.7e+02;
Matches 7; Conservative 4; Mismatches 9; Indels 2; Gaps 1;

QY 4 GAQDGRGVA--FNV 17
111 1111111
DB 7 GSEARGARGLASLFNM 22

RESULT 13

Q9QHT4 PRELIMINARY; PRT; 27 AA.
DI 01-MAY-2000 (1:EMBLrel, 13, Created)
DI 01-MAY-2000 (1:EMBLrel, 13, Last sequence update)
DI 01-JUN-2000 (1:EMBLrel, 14, Last annotation update)
DE POLYPROTEIN (FRAGMENT)
OS Hepatitis C virus.
OC Viruses; ssRNA positive strand viruses, no DNA stage; Flaviviridae;
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RA Sandres K., Dubois M., Pasquier C., Izopet J.,
RI "The genetic heterogeneity of hypervariable region 1 of the viral
RT genome and the sensitivity of hepatitis C virus to interferon alpha
therapy.";
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
RE EMBL: AF166843; AAF52501.1;
DR INTERPRO: IPR002541;
DR PFAM: PF01560; HCV_NSI; 1;
KW Polyprotein.
FT NON_TER 1 1
FT NON_TER 27 27
SQ SEQUENCE 27 AA: 2757 MW: 60AF66C4295BE22 CR064;

Query Match 29.9% Score 32; DB 14; Length 27;
Best Local Similarity 43.9% Pred. No. 1.7e+02;
Matches 7; Conservative 4; Mismatches 9; Indels 2; Gaps 1;

QY 4 GAQDGRGVA--FNV 17
111 1111111
DB 7 GSEARGARGLASLFNM 22

RESULT 14

Q9QHT3

